

From the Department of Microbiology, Tumor and Cell Biology
Karolinska Institutet, Stockholm, Sweden

REPERTOIRE DIVERSITY AND MATURATION OF HIV-1 VACCINE- INDUCED B CELL RESPONSES

Ganesh Eknath Phad



**Karolinska
Institutet**

Stockholm 2018

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.

Printed by Eprint AB 2018

© Ganesh Eknath Phad, 2018
ISBN 978-91-7831-086-9

REPERTOIRE DIVERSITY AND MATURATION OF HIV-1 VACCINE-INDUCED B CELL RESPONSES THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Ganesh Eknath Phad

Principal Supervisor:

Professor Gunilla Karlsson Hedestam
Karolinska Institutet
Department of Microbiology, Tumor and Cell
Biology

Co-supervisor(s):

Associate Professor Mats Persson
Karolinska Institutet
Department of Clinical Neuroscience

Assistant Professor Christopher Sundling
Karolinska Institutet
Department of Medicine, Solna

Opponent:

Professor Stephen Quake
Stanford University
Department of Bioengineering

Examination Board:

Professor Jan Albert
Karolinska Institutet
Department of Microbiology, Tumor and Cell
Biology

Associate Professor Annelie Tjernlund
Karolinska Institutet
Department of Medicine, Solna

Professor Nils Lycke
University of Gothenburg
Institute of Biomedicine
Department of Microbiology and Immunology

प्रिय आई आणि बाबांस

ABSTRACT

In the work presented in this thesis, I characterized B cell responses in rhesus macaques inoculated with HIV-1 envelope glycoprotein (Env)-based vaccines to define genetic and functional properties of the elicited antibodies for an improved understanding of Env immunogenicity.

In **Paper I**, we investigated antibody heavy chain V gene segment expression in the overall B cell repertoire and in response to vaccination by using 2 approaches, 454-pyrosequencing and single-cell RT-PCR of bulk and sorted memory B cells, respectively. Both methods gave highly concordant results and identified differential expression of the Ig gene segments. We further investigated the VH repertoire of antigen-specific memory B cells induced by an HIV-1 Env trimer immunogen. We found that the antibody response against Env was highly polyclonal consisting of most of the expressed VH segments in immunized subject. In **Paper II**, we isolated and comprehensively characterized genetic properties of a panel of vaccine-induced MAbs from Env immunized subjects. We demonstrated that the MAbs were genetically diverse with no bias in their V gene usage towards the distinct epitope regions they were mapped to, consistent with a highly polyclonal response. MAbs targeting two sub-regions, the CD4 binding site (CD4bs) and variable region 3 (V3), were neutralizing. The combined activity of these MAbs recapitulated the neutralizing activity observed in unfractionated plasma samples of the animals from which they were isolated, consistent with that these were the predominant specificities elicited. In **Paper III**, we used Illumina high-throughput sequencing (HTS) and developed a novel computational tool, IgDiscover, for germline Ig V allele identification. Previous work using targeted genomic PCR of germline V genes from Chinese rhesus macaques revealed a striking allelic diversity between animals and the presence of many novel alleles, highlighting the limitations of existing macaque V gene databases. By using IgDiscover to construct individualized Ig germline gene databases, we can improve the accuracy of Ab repertoire analyses in outbred populations considerably. In **Paper IV**, we investigated B cell responses induced by a new generation, well-ordered Env trimers. We used HTS sequencing to investigate how the response evolved over time and whether it disseminated into different immune compartments. We developed an IgDiscover extension module, *Clonoquery*, to trace hundreds of B cell clonal lineages targeting distinct sub-regions of Env in the peripheral blood, bone marrow, spleen, draining lymph node and gut. We detected broad dissemination of Env-specific B cell clones in all compartments measured, except gut. We showed that some lineages were greatly expanded and that the mean level of somatic hypermutation of the variants within each lineage increased with boosting, replacing previous variants from the same lineage in both blood and bone marrow.

In conclusion, this thesis offers a new information about the genetic and functional composition of Env-specific B cell responses in immunized rhesus macaques, a highly relevant model for understanding vaccine-induced response in humans. It also offers valuable information about the rhesus macaque Ig repertoire in general as well as broadly applicable tools for repertoire sequencing.

LIST OF PUBLICATIONS

- I. Christopher Sundling, Zhenhai Zhang, **Ganesh E. Phad**, Zizhang Sheng, Yimeng Wang, John R. Mascola, Yuxing Li, Richard T. Wyatt, Lawrence Shapiro and Gunilla B. Karlsson Hedestam. *Single-cell and deep sequencing of IgG-switched macaque B cells reveal a diverse Ig repertoire following immunization*. Journal of Immunology. 2014, 192(8):3637-44.
- II. **Ganesh E. Phad**, Néstor Vázquez Bernat, Yu Feng, Jidnyasa Ingale, Paola Andrea Martinez Murillo, Sijy O'Dell, Yuxing Li, John R. Mascola, Christopher Sundling, Richard T. Wyatt and Gunilla B. Karlsson Hedestam. *Diverse antibody genetic and recognition properties revealed following HIV-1 envelope glycoprotein immunization*. Journal of Immunology. 2015, 194(12):5903-14.
- III. Martin M. Corcoran*, **Ganesh E. Phad***, Néstor Vázquez Bernat, Christiane Stahl-Hennig, Noriyuki Sumida, Mats A.A. Persson, Marcel Martin, Gunilla B. Karlsson Hedestam. *Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity*. Nature Communications. 2016, 7:13642. *Shared first authors
- IV. **Ganesh E. Phad**[†], Pradeepa Pushparaj, Karen Tran, Paola Martinez-Murillo, Monika Adori, Viktoriya Dubrovskaya, Néstor Vázquez Bernat, Komal Bhullar, Sanjana Narang, Chiara Sorini, Eduardo Villablanca, Sijy O'Dell, John Mascola, Christopher Sundling, Marcel Martin, Ben Murrell, Martin Corcoran, Richard Wyatt and Gunilla B. Karlsson Hedestam[†]. *Clonal lineage tracing reveals dissemination and maturation of vaccine-induced B cell responses in multiple immune compartments* (Manuscript) [†]Corresponding authors

PUBLICATIONS NOT INCLUDED IN THIS THESIS

- I. Martinez-Murillo P*, Tran K*, Guenaga J, Lindgren G, Àdori M, Feng Y, **Phad GE**, Vázquez Bernat N, Bale S, Ingale J, Dubrovskaya V, O'Dell S, Pramanik L, Spångberg M, Corcoran M, Loré K, Mascola JR, Wyatt RT, Karlsson Hedestam GB. *Particulate Array of Well-Ordered HIV Clade C Env Trimers Elicits Neutralizing Antibodies that Display a Unique V2 Cap Approach*. Immunity. 2017, 46(5):804-817. *Shared first authors
- II. Wang Y, Sundling C, Wilson R, O'Dell S, Chen Y, Dai K, **Phad GE**, Zhu J, Xiao Y, Mascola JR, Karlsson Hedestam GB, Wyatt RT, Li Y. *High-Resolution Longitudinal Study of HIV-1 Env Vaccine-Elicited B Cell Responses to the Virus Primary Receptor Binding Site Reveals Affinity Maturation and Clonal Persistence*. Journal of Immunology. 2016, 196(9):3729-43.
- III. Madani N, Princiotta AM, Easterhoff D, Bradley T, Luo K, Williams WB, Liao HX, Moody MA, **Phad GE**, Vázquez Bernat N, Melillo B, Santra S, Smith AB 3rd, Karlsson Hedestam GB, Haynes B, Sodroski J. *Antibodies Elicited by Multiple Envelope Glycoprotein Immunogens in Primates Neutralize Primary Human Immunodeficiency Viruses (HIV-1) Sensitized by CD4-Mimetic Compounds*. Journal of Virology. 2016, 90(10):5031-46.
- IV. Dai K, Khan SN, Wang Y, He L, Guenaga J, Ingale J, Sundling C, O'Dell S, McKee K, **Phad GE**, Corcoran M, Wilson R, Mascola JR, Zhu J, Li Y, Karlsson Hedestam GB, Wyatt RT. *HIV-1 Vaccine-elicited Antibodies Reverted to Their Inferred Naive Germline Reveal Associations between Binding Affinity and in vivo Activation*. Scientific Reports. 2016, 6:20987.
- V. Klein F, Nogueira L, Nishimura Y, **Phad G**, West AP Jr, Halper-Stromberg A, Horwitz JA, Gazumyan A, Liu C, Eisenreich TR, Lehmann C, Fätkenheuer G, Williams C, Shingai M, Martin MA, Bjorkman PJ, Seaman MS, Zolla-Pazner S, Karlsson Hedestam GB, Nussenzweig MC. *Enhanced HIV-1 immunotherapy by commonly arising antibodies that target virus escape variants*. Journal of Experimental Medicine. 2014, 211(12):2361-72.
- VI. Navis M, Tran K, Bale S, **Phad GE**, Guenaga J, Wilson R, Soldemo M, McKee K, Sundling C, Mascola J, Li Y, Wyatt RT, Karlsson Hedestam GB. *HIV-1 receptor binding site-directed antibodies using a VHI-2 gene segment orthologue are activated by Env trimer immunization*. PLoS Pathology. 2014, 10(8):e1004337.
- VII. Sundling C, **Phad G**, Douagi I, Navis M, Karlsson Hedestam GB. *Isolation of antibody V(D)J sequences from single cell sorted rhesus macaque B cells*. Journal of Immunological Methods. 2012, 386(1-2):85-93.

TABLE OF CONTENTS

List of Abbreviations

1. Vaccines.....	1
2. B Cells and Antibodies.....	3
2.1 A Brief Overview of B Cell Development.....	3
2.2 Antibody Structure, genetics and diversity	3
2.3 Antigen-Specific B Cell Responses.....	8
3. Antibody Repertoires	11
3.1 Antibody Repertoire at the Single Cell Level	11
3.2 High-Throughput Sequencing (HTS) of Antibody Repertoires.....	11
3.3 Bioinformatic Analysis of Antibody Repertoire Data.....	14
4. Human Immunodeficiency Virus Type 1 (HIV-1).....	16
4.1 HIV-1 Structure	16
4.2 The Envelope Glycoproteins.....	16
4.3 Antibody Responses to HIV-1 Env In Infection.....	20
4.4 Env Trimer As A Template for HIV-1 Vaccine Design	21
4.5 Analysis of Env Trimer Vaccine-Induced B Cell Responses	23
5. Aims.....	26
6. Summary of Papers.....	27
6.1 Paper I	27
6.2 Paper II	28
6.3 Paper III.....	29
6.4 Paper IV.....	30
7. Concluding Remarks.....	31
8. Acknowledgements.....	32
9. References	38

LIST OF ABBREVIATIONS

Ab	Antibody
AID	Activation induced cytidine deaminase
AIDS	Acquired immunodeficiency syndrome
AIRR	Adaptive Immune Receptor Repertoire
ASC	Antibody-secreting cell
BCR	B cell receptor
BCMA	B cell maturation antigen
bNAbs	Broadly neutralizing antibodies
bp	Base pair (referring to the number of nucleotides)
CD4bs	CD4 receptor binding-site
CDR	Complementarity determining region
cryo-EM	Cryo-Electron Microscopy
CSR	Class-switch recombination
D	Diversity (region in immunoglobulin)
DZ	Dark zone (region in the germinal center)
ELISA	Enzyme-linked immunosorbent assay
ELISpot	Enzyme-linked immunospot
Env	HIV-1 Envelope glycoproteins
FACS	Fluorescence-activated cell sorting
FDC	Follicular dendritic cells
FR	Framework
Gag	Group-specific antigen
GC	Germinal center
gDNA	genomic DNA
HA	Influenza hemagglutinin
HIV	Human immunodeficiency virus
HTS	High-throughput DNA sequencing
ID	Inner domain
Ig	Immunoglobulin
IgH	Immunoglobulin heavy chain
Igκ	Immunoglobulin kappa chain
Igλ	Immunoglobulin lambda chain
IgL	Immunoglobulin light chain
IMGT	ImMunoGeneTics
J	Joining (region in immunoglobulin)
LLPC	Long-lived plasma cell

LZ	Light zone (region in the germinal center)
MAb	Monoclonal antibody
MBC	Memory B cell
MHC	Major histocompatibility complex
MPER	Membrane-proximal external region
Nef	Negative factor
NAbs	Neutralizing antibodies
OD	Outer domain
PBMC	Peripheral blood mononuclear cell
PC	Plasma cells
PCR	Polymerase chain reaction
Pol	Polymerase gene
pHSC	Pluripotent hematopoietic stem cells
RAG	Recombination activating genes
RSS	Recombination signal sequence
RT	Reverse transcriptase
SNP	Single nucleotide polymorphisms
SHM	Somatic hypermutation
TdT	terminal deoxynucleotidyl transferase
UMI	Unique molecular identifiers
V	Variable (region in immunoglobulin)
Vif	Viral infectivity factor

1 VACCINES

Vaccination is the most successful, powerful, and cost-effective medical measure that has saved more lives than any other health interventions existing to date. The world's first vaccine, the smallpox vaccine, was developed in 1796 by the pioneer Edward Jenner. His discovery has now led to the eradication of natural smallpox [1]. Almost a century after Jenner's breakthrough, Louis Pasteur introduced the process of attenuation and demonstrated its utility by producing a rabies vaccine derived from the killed extracts of infected central nervous system [2]. Since then, vaccines have evolved in many different types and forms [3-5] and the development continues in the 21st century with new strategies and techniques to produce efficacious vaccines that can protect the host against various infections, ensuring human longevity and health. At present, several licensed vaccines are available for the prevention of bacterial and viral diseases including the virus-induced cancers such as cervical cancer caused by human papilloma virus and liver cancer caused by hepatitis B virus [6, 7]. Nevertheless, there are several major global diseases such as AIDS, tuberculosis, and malaria that are responsible for considerable morbidity and mortality worldwide for which no licensed vaccines are available. The causative agents for these diseases either possess extensive genetic variability, intricate pathogenesis, or have evolved various immune evasion mechanisms, posing major challenges to vaccine development [6]. Moreover, most currently licensed vaccines have been developed empirically, with limited or no understanding of how they activate the immune system against the infection. It is now increasingly appreciated that empirical strategies alone are not sufficient for successful vaccine development against current pandemics such as those causing AIDS and malaria.

For most of the existing vaccines, antigen (Ag)-specific neutralizing antibodies (Ab) (as detected in vitro by ELISA, hemagglutination, or neutralization assays) have been shown to confer the protection against infections, and provide the best correlates of protective immunity [8]. Nevertheless, cellular immunity comprising of CD4+ T helper (Th) and CD8+ T cell responses are also important components of vaccine-induced protection against intracellular pathogens [9], especially in the elderly. Although the high specificity of antibodies is a key element of protective immunity, it is still largely unknown which sub-specificities are responsible for the protection against a given pathogen. Delineating the fine Ab specificities of experimental vaccines will provide important clues for designing improved immunogens against pathogens that easily escape immune recognition such as HIV-1.

Long-term maintenance of pathogen-specific antibody levels is crucial for establishing life-long protective immunity and is one of the most important attributes of successful vaccines. The durability of immune response can be determined by measuring the half-life of the polyclonal serum antibody response. The longitudinal maintenance of serum antibody responses against several vaccine antigens and infections has been determined where Ab half-life was shown to range between 11 years for non-replicating protein antigen, such as tetanus, and >3000 years for live attenuated measles virus [10, 11].

However, the mechanism(s) underlying the longevity of serum Ab response is not yet clear and remains the subject of debate. Three primary models/theories including, the polyclonal stimulation model [12], plasma cell niche competition/displacement model [12-14], and imprinted lifespan model of plasma cell longevity [15] have been proposed to explain the potential mechanisms contributing the maintenance of long-term protective immunity [15]. Multivalent antigens (e.g. measles virus particles) have been suggested to induce increased BCR cross-linking and Ag presentation to T follicular helper (T_{FH}) cells as compared to monovalent antigens (e.g. tetanus, diphtheria toxoids), resulting in genesis of long-lived Ab responses [15, 16]. The persistence of antigen, which is limited in case of subunit vaccines compared to natural infection or immunization with replicating viral vectors, could be an additional critical factor for the generation of durable, high-titer Ab responses and thus, long-lasting immunity.

2 B CELLS AND ANTIBODIES

2.1 A BRIEF OVERVIEW OF B CELL DEVELOPMENT

B lymphocytes are central in the humoral immune system and they are produced continuously throughout life. The development of B cells begins in the fetal liver and continues in the bone marrow, differentiating from pluripotent hematopoietic stem cells (pHSCs) to immature B cells in a number of distinct steps. During differentiation in the bone marrow, stepwise gene rearrangements at the immunoglobulin (Ig) heavy and light chain loci in B cells result in the generation of functional B cell receptors (BCRs) [17]. In the bone marrow microenvironment, immature B cells expressing functional BCRs are subjected to negative selection by self-antigens in a process called central tolerance. During this process, B cells expressing self-reactive BCRs undergo either clonal deletion or receptor editing or become anergic, whereas B cells with no autoreactivity enter into the circulation. Immature B cells with non-self-reactive BCRs after passing this central tolerance checkpoint migrate to the spleen, lymph nodes or gut-associated lymphoid tissues and continue to differentiate into mature naïve B cells expressing both IgM and IgD BCRs. During the maturation, residual B cells expressing autoreactive BCRs are eliminated in a process known as peripheral tolerance [18]. B cells passing both central and peripheral checkpoints generate the naïve B cell repertoire that is mostly devoid of self-reactive clones and capable of recognizing and neutralizing a plethora of non-self-antigens and pathogens.

2.2 ANTIBODY STRUCTURE, GENETICS AND DIVERSITY

Antibodies are heterodimeric protein complexes consisting of pairs of disulfide-linked immunoglobulin heavy (IgH) and light (IgL) chains (Figure 1). Each heavy chain (μ , α , γ , δ , ϵ) and light chain (κ , λ) contains functionally distinct variable and constant domains, present at the N- and C-terminal regions, respectively, each consisting of approximately 110-130 amino acids [19]. The variable domains are composed of the complementarity-determining regions (CDR1-3) as well as the adjacent framework regions (FR1-4), which together form a variety of antigen-binding tertiary structures [20]. The variability in the variable domains is concentrated in the CDR loops in which the heavy chain CDR3 is the most diverse region due to the stochastic recombination and non-templated nucleotide additions, while the other CDRs are subjected to variation during the process of SHM (**Paper IV**). Usually, the CDR3 is the principle determinant of antibody specificity [21, 22]. However, the CDR1 and CDR2 regions can also contribute significantly to antigen binding and in some cases antigen specificity is determined predominantly or entirely by the light chains. The framework regions, present as β -pleated sheets, are less variable and primarily act as scaffolds that support the hypervariable CDR loops. The constant regions of the heavy chain (μ , δ , γ 3, γ 1, α 1, γ 2, γ 4, ϵ , and α 2) encode the constant domains, which determine the class and effector function of antibody molecules [19].

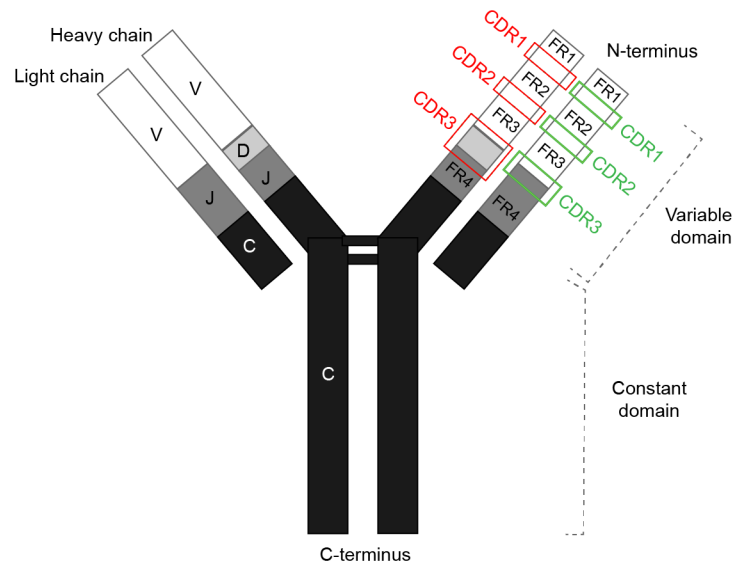


Figure 1: Schematic of an antibody molecule. The CDR1-3 regions are shown in red and green boxes for heavy and light chains, respectively.

2.2.1 The immunoglobulin (Ig) loci

The immunoglobulin loci encode the many gene segments that recombine to form the template for heavy and light chains of Ab molecules. The heavy chain is expressed from the IgH locus and the light chain from one of two light chain loci, the lambda (Ig λ) or kappa (Ig κ) locus. Matsuda *et al.*, determined the first full-length physical map of the human IgH locus, a 957kb region containing multiple variable (V), diversity (D), and joining (J) gene segments located on the distal end of chromosome 14 (14q32.33). DNA encompassing the entire IgH locus revealed that it consists of a highly interspersed organization of functional and non-functional (pseudogenes) V gene segments belonging to seven different families in addition to clusters of D and J germline gene segments [23]. Later, the Ig κ locus comprising V κ genes that are organized in two clusters (proximal and distal) and J κ segments, covering 1,820 kb, was mapped to chromosome 2 (2p11.2) [24-26], and the Ig λ locus consisting of V λ and J λ genes was mapped to chromosome 22 (22q11.2) spanning 1,050 kb [27-29]. In both IgH and Ig κ loci, the constant (C) region is encoded by separate exon(s) that are located downstream of the J gene segments, while for Ig λ , there are several C regions interspersed between the J gene segments (reviewed in [30]). In rhesus macaques, which are estimated to share ~93% homology to the human genome, the overall organization of the Ig loci is similar to those of the human Ig loci, but the V, D (heavy only), J and C genes encoding heavy, kappa, and lambda chains are located on chromosomes 7, 13 and 10, respectively in macaques [31-33].

The presence of multiple V, D, and J germline gene segments and their recombination is critical to the generation of a diverse Ab repertoire. Despite their importance, our understanding of structural and allelic variation in germline genes between individuals is far from complete [34, 35]. This is mainly because of the highly complex nature of the Ig loci. Genomic analyses of the Ig loci have demonstrated that it comprises multiple

repetitive elements and many highly similar genes considered to have evolved via gene duplication, divergence [36], and conversion [37]. Due to this complex genomic architecture, comprehensive sequencing, assembly and annotation of the Ig loci is difficult to achieve [34]. In addition, the Ig loci, specifically IgH, is highly polymorphic [23]. A number of early and recent studies using different approaches have shown a variety of polymorphisms, including insertion/deletion of VH segments and single nucleotide polymorphisms (SNPs) in the IgHV region. The detection of VH polymorphisms has resulted in the identification of several polymorphic VH alleles [38-47], contributing significantly to the VH gene diversity in the population. For humans, a number of V gene alleles within the IgH, Igk, and Igλ loci are cataloged in the ImMunoGeneTics (IMGT) database [48-52]. However, the IMGT database significantly lacks coverage of the full allelic diversity in the human population. For example, sequencing of samples from different ethnic groups revealed a number of population-specific novel alleles [47, 53]. There is currently growing interest in understanding allelic variation across ethnically diverse populations, as this would help to interrogate the impact of Ig germline gene polymorphisms on antibody response in infection, vaccination, and autoimmune diseases [54]. However, so far, the full nucleotide sequence of ~1Mb IgH assembly (V, D, and J gene region except C region) has only been determined only twice [23, 55], limiting our knowledge about genetic variation in the Ig loci across the population. So far, the size and complexity of the Ig loci have hindered genomic sequencing of complete Ig loci from large cohorts of individuals, as it requires both longer reads and higher levels of coverage for efficient sequence assembly. In addition, Ig alleles obtained from genomic sequencing encompass both functional and non-functional/pseudogenic alleles, which are often difficult to distinguish from each other, posing an additional challenge for genomic mapping of Ig gene segments [34]. While comprehensive genomic sequencing of the full Ig loci remains an obstacle, cDNA-derived expressed V(D)J sequences are now increasingly seen as a useful resource to explore the Ig gene allelic variation among individuals. Novel computational approaches and tools have been developed and applied to analyze expressed V(D)J sequences. Application of such germline gene inference tools, including IgDiscover (**Paper III**) and TIgGER [56], has led to the identification of several novel V germline gene alleles that are not present in current germline database resources. Thorough validation of newly identified alleles using these tools is important and include targeted genomic sequencing, detection of the same allele in multiple individuals [47, 57] or haplotype inference methods [58]. *De novo* identification of germline alleles from Ab-encoding V(D)J transcripts using tools such as IgDiscover enables the construction of personalized germline databases without any prior knowledge of an individual's genetic background. The same approach can be applied to create databases of germline alleles for species for which complete reference genomes are lacking. Using this approach, we generated individualized germline gene databases for a number of rhesus macaques by applying IgDiscover, which revealed that the allelic diversity between the animals was extensive [59]. Thus, when correctly used, computational approaches can be far more efficient than genomic sequencing for studies of antibody germline gene variation at the population level. Knowledge of an individual's germline alleles, or access to a comprehensive database of functional germline alleles, is important for studies of B cell responses since

it enables correct germline gene assignment and accurate calculation of somatic hypermutation [54].

2.2.1.1 Rearrangement of antibody V(D)J segments

Rearrangement of the Ab heavy and light chain occurs at different stages of B cell development. Recombination of the heavy chain genes takes place before the light chain gene assembly begins. The germline sequences encoding each V gene segment of the heavy or light chains are comprised of two small upstream exons separated by a short intron (Figure 2). An N-terminal signal sequence (or leader sequence) directs the translated Ab polypeptide through the rough endoplasmic reticulum which is later cleaved from the nascent heavy or light chain in the post-translational processing.

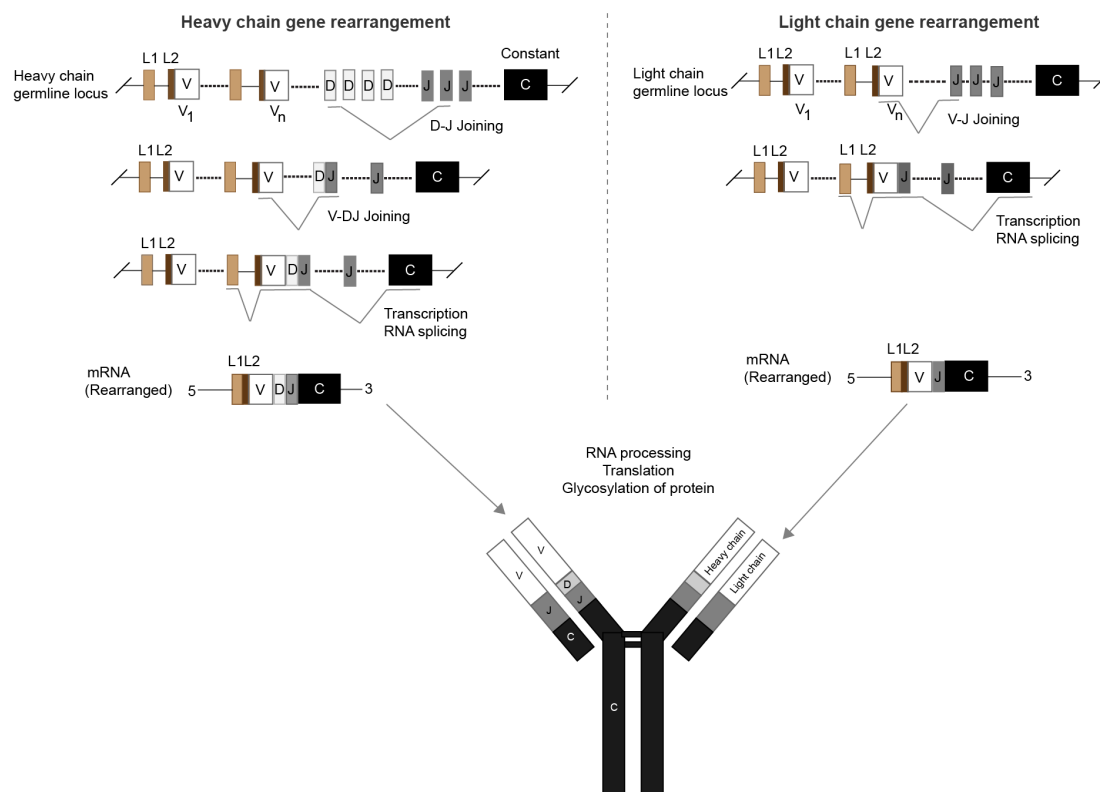


Figure 2: Schematic of antibody heavy chain and light chain gene rearrangement process (V-(D)-J recombination) Leader 1 (L1) and leader 2 (L2) shown in light and brown colors, respectively.

In the genomic configuration of the Ig loci, the V, D, J and C gene segments are present as discrete clusters. During B cell development, these V, D, and J gene segments undergo somatic gene rearrangement events to form unique variable region exons. Each functional V, D, and J gene segment is flanked by recombination signal sequences (RSSs). The RSSs consist of highly conserved heptamer and nonamer sequences separated by either 12 or 23 base pair (bp) spacers, respectively corresponding to one and two turns of the DNA helix and referred as one-turn, or the 12RSS, and two-turn, or the 23RSS. Effective somatic recombination occurs only between 12RSS and 23RSS flanked gene segments. For instance, in the IgH locus each V and J gene segment is

flanked by 23RSSs, whereas each D segment is associated with 12RSS. This prevents direct V to J joining and allows efficient recombination between, initially the D and J segment, followed by joining between the V segment and newly generated D-J product. The resultant recombined V(D)J is subsequently spliced to the downstream constant region, which typically is the μ -domain, leading to the synthesis of IgM BCRs in developing naïve B cells. This stepwise rearrangement process is catalyzed by several enzymes, collectively termed the V(D)J recombinase, which are involved in double-stranded DNA break and repair. During recombination, the ends of the gene segments are prone to imprecise exonuclease digestion, and non-germline encoded random bases (N bases) are added at the heavy chain VD, DJ, and light chain VJ junctions by the terminal deoxynucleotidyl transferase (TdT) to generate the additional diversity at the V(D)J junction that encodes the HCDR3 loop. The detailed mechanisms and regulation of the V(D)J recombination events have been reviewed elsewhere [19, 60-62]. Following the somatic recombination, only productively rearranged heavy and light chains are assembled as antibody heterodimers that can either be expressed as membrane-bound BCRs or secreted in the form of soluble antibodies [19].

2.2.2 Antibody Diversification

Prior to exposure to antigen, the generation of a diverse antibody repertoire within a given individual occurs early during B cell development in the bone marrow. From combinatorial assembly of V, (D), and J segments, junctional diversity generated due to imprecision of the recombination process, and combinatorial pairing of IgH and IgL chains. Different individuals encode different germline V(D)J alleles and it is also becoming increasingly clear that there is significant structural variation in the Ig loci with frequent gene segment copy number variation, adding further diversity in the population [35]. Additional rare processes such as receptor editing and replacement of V gene segment can contribute further mechanisms for antibody repertoire diversity. It is estimated that these mechanisms can generate a primary repertoire of $>10^{11}$ different antibodies [19, 60]. This enormous diversity of the primary naïve B cell repertoire is further augmented after antigenic stimulation by two mechanisms - somatic hypermutation (SHM) and class-switch recombination (CSR), leading to Ag-dependent secondary diversification of the Ab repertoire. Both SHM and CSR increase the efficacy of active immune response and are controlled by the activation-induced (cytidine) deaminase (AID), an enzyme expressed in the germinal center (GC) B cells. A recent study by Shlomchik and coworkers showed that SHM is not only confined to the GC but can also occur at extrafollicular sites, leading to affinity matured Ab response against *Salmonella*. Their results suggest that functionally active AID is expressed in non-GC B cells during the EF response [63].

2.2.2.1 Somatic Hypermutation (SHM) and Class-Switch Recombination (CSR)

After immunization or infection, activated mature B cells expressing surface IgM with low-to-moderate affinity for their cognate antigen migrate to the GCs in secondary lymphoid organs (SLO), where they mutate their rearranged V(D)J sequences via SHM as well as alter the C region of the heavy chain via CSR. AID-mediated SHM preferentially targets the RGYW or its complement, WRCY (where R=A or G, Y=C or T, W=A or T) mutational “hotspot” motifs [64-68]. SHM diversifies the Ab repertoire by introducing non-templated point mutations at a rate of 10^{-5} – 10^{-3} /base pair/cell division in the Ig genes, generating clonal offspring resulting in antibodies that may acquire enhanced binding affinity for a given antigen. The mutations are confined to the V(D)J coding region, ~100–200 bp downstream of the transcription initiation site and exponentially decaying at the 3' end, approximately 1.5–2 kb downstream. Thereby, SHM is avoided in the critical regulatory elements in the promoter, the intronic enhancer, and the constant (C) region genes, the latter which is responsible for the antibody effector functions [69, 70]. The CSR, also termed as isotype switching, on the other hand involves deletional DNA recombination that occurs between the switch regions to replace the C-region ($C\mu$ and $C\delta$) exons with a different set of downstream IgH C-region exons, such as $C\gamma$, $C\alpha$ or $C\epsilon$. For example, activated naïve B cell can switch from expressing IgM and IgD on their surface to expressing IgG, IgE, or IgA. This class switching allows Abs to have different effector functions while maintaining the same antigen specificity. Although less common, CSR can occur prior to GC formation [71] and also during T cell-independent responses, which do not induce GCs [72].

2.3 ANTIGEN-SPECIFIC B CELL RESPONSES

2.3.1 B cell Activation and Germinal Center reaction

Following antigen exposure, mature naïve B cells can recognize and capture their cognate antigen in its native form via their BCRs. After priming with the cognate antigen, activated Ag-specific B cells internalize, process, and present antigenic peptide fragments on major histocompatibility complex (MHC) class II molecules, and begin to up-regulate the CCR7 and EBI-2 receptors, which promote the migration of B cells to the B cell - T cell borders of the draining lymph node, where they can receive help from cognate T helper cells [73-75]. At the B - T cell interface, Ag-primed B cells begin to proliferate upon receiving signals from T cells. After 1-2 days, proliferating B cells develop along two pathways; in the first pathway, most B cells with relatively high affinity BCRs differentiate into short-lived extrafollicular plasma cells (PC), which provide an immediate Ab response to antigen; while in the second pathway, lower-affinity B cells, in addition to the remainder of the high affinity B cells, enter the GC reaction for rounds of SHM and selection, a process in which the cells that acquire mutations leading to reduced affinity die [76]. Currently, the mechanisms and signals that promote B cell differentiation along these two pathways remain incompletely defined.

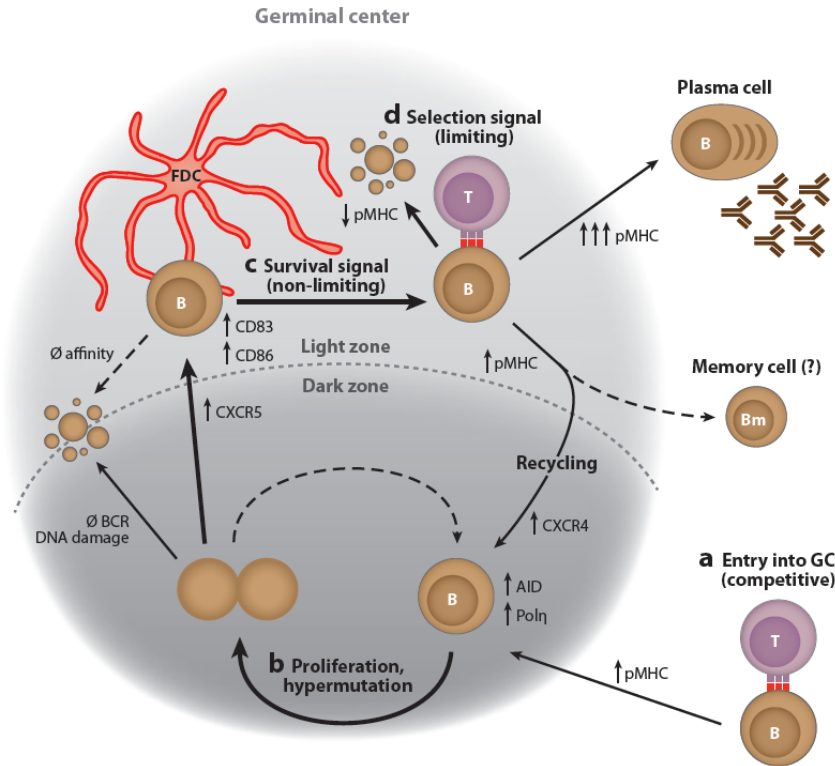


Figure 3: Schematic of proposed model of germinal center (GC) reaction (adapted from [77]). At the B cell-T cell border of the SLOs, antigen primed B cells process and presents the antigen to T helper cells and competes to receive the co-stimulatory signals (a). The selected B cells are allowed to enter the dark zone (DZ) of GC where they proliferate and mutate their cell surface receptors (BCRs) via SHM process by up-regulating the components of SHM machinery including AID and the Pol η enzyme (b). After one or possibly several rounds of proliferation and mutation in the DZ, the B cells migrate to the light zone (LZ) of GC where they test their mutated BCRs against the antigen present in the form of immune complexes on follicular dendritic cells (FDCs) (c). B cells with very low or no affinity for antigen undergo apoptosis due to lack of survival signals, while remaining B cells need to compete for selection signal from limited T helper cells (d). B cells with higher affinity will more successfully compete for T cell help and lower affinity B cells will undergo apoptosis. After successful interaction with T helper cells, surviving B cells have three potential fates: re-entry in to the DZ for further division/mutation, exit the GC as plasma cells or exit the GC as memory B cells.

GCs are transient structures that appear within days in the B cell follicles of SLOs in response to T cell-dependent antigens. B cells participating in the GC reaction selectively up-regulate the transcription factor Bcl-6, which plays critical roles in supporting the survival of GC B cells and preventing premature exit from the GC [78]. In addition, depending on the expression of CXCR4, GC B cells transit between the dark zone (DZ), where they undergo extensive proliferation and SHM, and the light zone (LZ), where they are exposed to antigen presented by follicular dendritic cells (FDCs). After several rounds of proliferation and SHM, B cells from the DZ migrate to the LZ where they test their mutated BCRs for binding to Ag presented on the FDCs. In the LZ, B cells with very low affinity BCR will not receive the survival signal and will undergo apoptosis [79], whereas the remaining B cells compete for limited T cell help during which the high affinity clones outcompete the low affinity clones [80, 81]. Surviving B cells then

either re-enter the DZ for further SHM or exit the GC as long-lived Ab-secreting plasma cells (LLPC) or as non-secreting memory B cells (MBC) (Figure 3).

2.3.2 Memory B Cells and long-lived plasma cells

Following the GC reaction, selected high affinity B cells differentiate into quiescent MBCs or LLPCs, which together contribute to B cell memory and thus, humoral immunity. The signals that determine whether a GC B cell differentiates into MBC or LLPC remains poorly understood but asymmetric cell division has been suggested [82, 83].

MBCs circulate through the SLOs and the peripheral blood and can also be found in the splenic marginal zone [84, 85], mucosal epithelium [86], or near contracting GCs within SLOs [87] where they persist for long periods of time in the absence of cognate antigen or T cell help [88-90]. Following antigen re-challenge, MBCs proliferate and rapidly differentiate into short-lived Ab-secreting cells (ASC) [91, 92], aiding to clear the infection, together with the Ab response generated during first encounter with cognate antigen arising from LLPCs. While it is still not clear whether MBCs can re-enter the GCs upon Ag re-encounter, it has been proposed that IgM⁺ MBCs can re-gain GC entry whereas GC derived IgG⁺ MBCs differentiate into ASCs [93, 94].

LLPCs exit from the GC as terminally differentiated ASCs that secrete large amounts of high-affinity Ag-specific Abs, providing a first line of defense against recurring infections [10, 15, 95]. The majority of the LLPCs are bone marrow (BM) resident where they are dependent on survival signals/factors such as APRIL, CXCL12, BAFF and IL-6 produced by surrounding cells including stromal cells, eosinophils, basophils, and megakaryocytes. These cells collectively form the plasma cell survival niche, which is responsible for viability of the plasma cells in the BM. The homing of LLPCs to a survival niche is dependent on the expression of surface chemokine receptors. For example, upregulation of CXCR4 receptor expression leads the migration of LLPCs to the bone marrow [96, 97], CCR5 and CCR28 to the mucosa [73], and CXCR3 to sites of inflammation [73]. In the BM, survival signals such as APRIL and BAFF mainly act via B cell maturation antigen (BCMA) receptor by inducing a critical anti-apoptotic protein, Mcl-1, which is essential for survival of BM plasma cells [98]. BM-resident LLPCs are maintained independently of memory B cells [99, 100] and can persist for decades in the absence of antigen [101, 102]. While LLPCs are believed to be the main source of circulating IgG Abs and it has been shown that serum Ab titers derived from LLPCs in response to antigens and viruses have a wide range of half-lives (10 to > 3000 years) [10]. Recently, Lee and colleagues characterized the BM plasma cells in humans where the CD19⁻ CD38^{hi} CD138⁺ subset, but not the CD19⁺ CD38^{hi} CD138⁺ subset was identified as LLPCs using morphological, RNA transcriptome, and heavy chain (VH) repertoire analyses [103]. However, further studies are required to validate these studies using higher resolution methods and functional evaluation of antigen-specific LLPCs.

3 ANTIBODY REPERTOIRES

3.1 ANTIBODY REPERTOIRE AT THE SINGLE CELL LEVEL

Recent developments in single-cell RT-PCR coupled with Sanger sequencing, cloning of Ig genes and expression of recombinant MAbs in mammalian cells allow the examination of Abs encoded by up to hundreds of single B cells with distinct specificities. The isolation of MAbs is a powerful approach that has improved our understanding of humoral immune response in infections [104], vaccinations [105, 106], and autoimmune diseases [107, 108] or immunodeficiency syndromes [109, 110] at the molecular level. In addition, many MAbs have therapeutic potential and their use represent an important strategy for immunotherapy [111]. A plethora of potent broadly neutralizing antibodies (bNAbs) against clinical pathogens including HIV-1 [112], influenza [113], malaria [114], dengue [115], and Zika [116] have been isolated. Detailed characterization of binding epitopes and neutralization mechanisms of some of these infection-induced MAbs have revealed sites of vulnerability on pathogens such as HIV-1 and an increased understanding of the evolution of neutralizing Abs during infection [117, 118]. This information has led researchers to design, develop, and test several structure-based vaccine candidates in relevant animal models with the aim of eliciting similar neutralizing antibody responses [54, 119-121]. Furthermore, the isolation and in-depth analyses of MAbs elicited by such vaccine candidates have helped to improve our understanding of vaccine-induced B cell responses at the genetic, functional [32, 121-123], and structural level [124]. As shown in this thesis and elsewhere, Sanger sequencing of hundreds to thousands of BCRs of vaccine-elicited single B cells have revealed the genetic diversity of Ab responses towards specific epitope or antigen (**Paper I**) [125]. In addition, identification of antigen-specific BCR sequences allows tracing of individual B cell clones over time, which helps our understanding of how they distribute across different immune compartments and how they evolve over time (**Paper IV**). High-resolution studies of vaccine-induced antibody responses has important implications for the re-design of existing immunogens to improve the elicited responses.

3.2 HIGH-THROUGHPUT SEQUENCING (HTS) OF ANTIBODY REPERTOIRES

Single B cell sequencing and cloning methods have provided invaluable information; however, they can only characterize 10^2 – 10^3 B cells per experiment. Therefore, using this approach alone is unrealistic to capture the astonishing diversity of B cell repertoires considering the labor and cost required for such endeavors. In contrast, high-throughput DNA sequencing (HTS) of Ab repertoires allow survey of thousands to millions of different clones with unique Ab gene rearrangements, providing insight into Ab repertoires at an unprecedented level of detail.

3.2.1 Experimental considerations

HTS technologies have developed rapidly over the past decade [126]. However, they are still evolving in capacity, capability, and applications and a number of challenges remain to unravel their full potential. Compared to Sanger sequencing read length and accuracy, common HTS platforms used for Ab repertoire studies produce relatively short reads and have higher error rate, which together with potential bias introduced during library preparation can skew conclusions drawn from the data. Therefore, sequencing Ab repertoires requires careful experimental design, with the major considerations being: 1) choice of sequencing platform 2) source and population of B cells 3) choice of template (genomic DNA or cDNA), 4) choice of primers in amplifying V(D)J rearrangements, gene-specific primer mix (multiplex PCR) versus unbiased methods such as 5' rapid amplification of cDNA ends (5'RACE) and addition of unique molecular identifiers (UMI) during cDNA and/or PCR amplification steps and 5) robust computational platforms for analysis of the data.

Sequencing platform:

The most commonly used HTS platforms to interrogate the Ab repertoires are the Illumina MiSeq and the Roche 454, each of which offers different advantages and disadvantages. The majority of earlier repertoire sequencing studies used the 454 platform, but in recent years, the development and support to the 454 platform was discontinued and eclipsed by a better and more precise low cost sequencing technology, the Illumina MiSeq platform. Currently, the Illumina MiSeq with its improved read length, relatively low error rates, and lower costs per base has become a dominant and preferred HTS platform in most of the Ab repertoire investigations [127-129]. Other platforms such as PacBio and IonTorrent have been used in repertoire sequencing [130, 131], but are less optimal due to either lower throughput and/or higher sequencing error rate [132].

Source and population of B cells:

In humans, the most common and easily accessible compartment for B cells is the peripheral blood, which used as a primary sample source in many Ab repertoire studies. However, studies have estimated that only 2% of total B cells are present in the peripheral blood, while remaining B cell diversity distribute between the lymph nodes (28%), the spleen and mucosal surfaces (23%), and the red bone marrow (17%) [133]. Therefore, sequencing Ab repertoires from peripheral blood does not provide a comprehensive picture of B cell diversity in the human body. For ethical reasons, access to tissue compartments other than the blood is limited for humans compared to animal models such as rhesus macaque or mouse. Rhesus macaques represent an ideal model for Ab repertoire studies in compartments that are inaccessible in humans (**Paper IV**). In comparison, mice are different than humans not only in their antibody genetics [134] but also in the microanatomy of their tissues [135] and in the definition of their B cell subsets [136].

Choice of template (genomic DNA or cDNA):

The DNA template for the construction of libraries for Ab repertoire sequencing can be isolated from total lymphocytes, peripheral blood mononuclear cells (PBMCs) or any tissue compartment containing B cells. The template may be isolated from a specific B cell subset, such as plasmablasts, plasma cells or memory B cells enriched for by flow cytometry or magnetic bead separation. Depending on the aim of study, genomic DNA (gDNA) or mRNA-derived cDNA can be used as a template to prepare the Ab libraries. The use of gDNA allows estimation of the proportion of unique B cell clones in the cell populations because of the presence of a single rearranged BCR per cell. Quantification of cell abundance is more challenging with mRNA as a template because of non-uniformity of Ab gene expression in different subsets of B cells. However, an important advantage of mRNA as a starting material is that it enables the identification of isotype information, which is not possible with gDNA because of the long intron between the J gene and constant region [137, 138].

Multiplex or 5'RACE PCR:

The amplification of rearranged V(D)J sequences is typically performed with multiplex PCR using a mixture of primers targeting the 5' region of multiple V genes, together with an isotype-specific heavy or light chain primer. Alternatively, the 5'RACE approach may be used, which enables the amplification of V(D)J sequences with single primer instead of primer mix can be used to minimize the potential amplification bias towards specific V gene segments [137]. Although the 5'RACE method is thought to result in less biased libraries, preferential and biased amplification was observed in TCR sequencing [139], possibly due to variable efficiency of the template switching reaction. Also, the 5'RACE protocol is generally less efficient than 5'Multiplex [140] and it is limited to RNA as a starting material. In addition, 5'RACE is less robust when used on RNA of suboptimal quality or limited quantity resulting in high non-specific amplification and short fragment contamination from RNA degradation or incomplete cDNA synthesis and template switching [141-143]. Furthermore, 5'RACE generates longer PCR products (due to amplification of 5' untranslated regions), which presents a significant challenge to current sequencing technology (usually Illumina MiSeq) used in repertoire sequencing. An advantages with 5'RACE is that it enables the identification of 5' untranslated regions (UTR) and upstream leader sequences of V alleles, information that can be used when designing 5' multiplex primers for species where the genomic information is limited [59]. Prior to PCR amplification, nucleotide barcodes (also called as unique molecular identifiers), which are random or semi-random short nucleotide sequences can be appended to each mRNA template during RACE or gene-specific cDNA synthesis at 5' or 3', respectively. Both multiplex PCR and RACE protocols have the potential to affect the quantitative aspects of Ab repertoire due to the possibility of heterogeneity in amplification efficiency. Therefore, molecular barcoding is essential as it allows accurate quantification of the DNA template in the sequencing library and correction of sequencing errors (by building consensus of multiple reads having the same barcode), enabling error-free immune profiling [137].

3.3 BIOINFORMATIC ANALYSIS OF ANTIBODY REPERTOIRE DATA

High-throughput sequencing of Ab repertoires generates vast amounts of data that contain immunological information of enormous complexity. Rigorous analyses of these datasets are required to filter and parse the information into interpretable results, but once this is achieved the results can greatly improve our understanding of adaptive immunity in health and disease [144]. Understanding the formation and dynamics of adaptive immune responses during infection, vaccination, malignancy, and autoimmune diseases is of particular relevance for the development of novel immunotherapies, vaccine candidates and immunodiagnostics. Therefore, there is substantial interest in this field, which, in turn, has led to the development of a variety of sophisticated computational algorithms and bioinformatic workflows for interrogation of Ab repertoires in humans and other species [138, 144]. Ab repertoire analysis is a multistep process and the key steps include pre-processing (error correction, read filtering) of sequencing data followed by V(D)J germline annotation, and clonotyping/clonal grouping. The output obtained from the latter steps can be processed further for secondary quantitative and qualitative analyses, such as clonal diversity estimation, selection estimation and analysis of SHM, clonal evolution and expansion, and clonal tracing, using numerous methods available for such purposes [138]. The wide variety of computational approaches means that there are no single standard operating procedure or guidelines, either experimental or computational, available for the Ab repertoire analysis. Although it remains a challenge to standardize the experimental methods or data processing pipelines due to continually evolving sequencing methods and computation tools, considerable efforts are being taken by the Adaptive Immune Receptor Repertoire (AIRR) Community to develop minimum standards and recommendations for repertoire sequencing studies (<http://www.airr-community.org>) [145, 146].

Preprocessing:

As discussed above, HTS of Ab repertoires suffer from errors and biases introduced during library preparation (PCR artifacts, amplification bias) and during the sequencing reaction (platform-specific). Therefore, preprocessing of HTS data is mandatory to address these technical issues. The purpose of the preprocessing step is to obtain error-corrected Ab sequences from raw reads produced by a sequencing instrument. Since the biological interpretation of the analysis directly depends on the quality of the data, several experimental and computational approaches have been developed and combined together to mitigate the impact of errors on the conclusions. The errors and bias in repertoire sequence data can be corrected through UMIs, replicate sequencing, sequence clustering, quantitative bulk PCR, and via Phred quality score-based filtering [138, 147, 148].

V(D)J germline gene segment annotation:

Accurate V(D)J germline gene assignment of error-corrected Ab sequences is critical for Ab repertoire analysis depends on how comprehensive and accurate available databases are. Several web-based or stand-alone annotation tools of variable efficiency and

accuracy such as IMGT [50], IgBLAST [149], iHMMune-align [150], and MIXCR [151] that are based on different algorithms, can be used for V, D, and J gene assignments. In general, assignment of the V and J gene segments is more reliable than D gene assignment because D genes typically are very short and sometime undergo trimming during the rearrangement process, which reduces their length further, making accurate assignment difficult [152]. The database of human and mouse germline V, D, and J sequences accessible through the IMGT repository has been used in almost all published Ab repertoire studies so far and as such has been tremendously useful. However, through recent HTS work, it has become clear that the allele collection within the IMGT repository is incomplete and does not cover the full extent of human genetic variation, specifically in V genes [35]. Thus, an important goal is to generate a publicly available, comprehensive database of germline V(D)J genes is required for accurate analysis of BCR repertoires, an effort that is ongoing in our laboratory in collaboration with the AIRR community. In addition, comprehensive germline gene databases for animal models frequently used in biomedical research such as rhesus macaques, guinea pig, and rabbit do not exist, currently limiting studies of Ab responses in these species. Germline gene assignment using any database that is not complete or under-representative of allelic diversity in a given species will affect the accuracy of clonal lineage tracing and SHM estimations and therefore impact immunological interpretations [54]. The IgDiscover tool, developed in **paper IV**, will be extremely helpful in the endeavor to establish accurate and more complete databases of antibody germline genes for humans and animals, which in turn will improve the overall of quality of Ab repertoire analysis going forward.

Clonotyping (clonal grouping):

After V(D)J annotations, Ab sequences based on their genetic information can be grouped into B cell clones, a process commonly referred to as clonotyping or clonal grouping. B cell clones that are descended from a common ancestor usually differ in their sequences due to SHM in their Ab-encoding genes, therefore, defining clonal relatedness in large sets of mutated Ab sequences is challenging [153]. A common method used currently for clonal grouping is to cluster all sequences with the same germline V, J allele assignment and CDR3 length along with some level of similarity between their CDR3 regions. The CDR3 region formed during V(D)J rearrangement is defined for each B cell clone and usually remains constant during the affinity maturation process. CDR3 length is therefore considered a primary indicator for clonal relatedness. Clonal grouping by clustering CDR3s at nucleotide [106, 154] or amino acid level based on 80-100% CDR3 homology have been performed [148, 155-157]. Different identity thresholds impact the clonal grouping differently (**paper IV**), potentially affecting downstream analyses and biological conclusions drawn from them. Currently there are no standard criteria or robust definition available for the accurate clonal grouping.

4 HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 (HIV-1)

4.1 HIV-1 STRUCTURE

The human immunodeficiency virus-1 (HIV-1) is the causative agent of acquired immunodeficiency syndrome (AIDS), a global pandemic with over 35 million cases of infection and 1.5 million annual deaths worldwide [158]. HIV-1 belongs to the lentivirus genus of the *Retroviridae* family. The members of the *Retroviridae* family contain two positive-sense RNA strands enclosed in a cone-shaped capsid, which in turn is contained within the icosahedral structure of the assembled virus particle. Upon budding, the virus acquires a lipid envelope derived from the host cell membrane. Upon infection of a new target cell, the ~9.2 Kbp long RNA molecules are reverse transcribed into double stranded DNA by the extremely error-prone, virally encoded enzyme, reverse transcriptase (RT). This process generates the enormous strain diversity of HIV-1 by introducing random point mutations in the viral genome. Like other retroviruses, the HIV-1 genome encodes three major genes, 5'-*gag-pol-env*-3'.

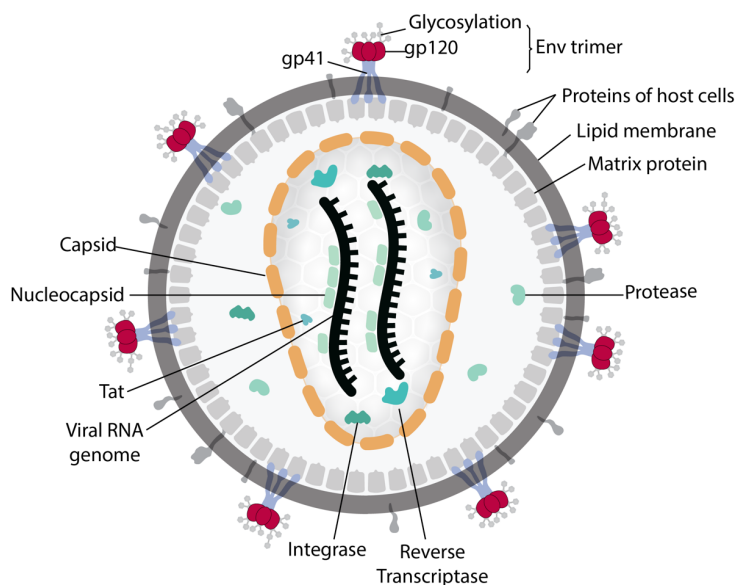


Figure 4: Schematic of HIV-1 virus (modified Thomas Splettstoesser, www.scistyle.com)

The *gag* (group-specific antigen) gene-encoded polyprotein is processed by the viral protease to produce the matrix (p17), capsid (p24), nucleocapsid (p7), spacer peptides (p1 and p2), and p6 proteins. The *pol* (polymerase) gene codes for the three viral enzymes; protease, reverse transcriptase, and integrase that are essential for the viral life cycle. The *env* (envelope) gene encodes the gp160 envelope glycoprotein precursor, which is extensively modified during biosynthesis as described below. The HIV-1 genome also contains the genes encoding for the accessory proteins (*vif*, *vpr*, *vpu*) and the regulatory proteins (*tat*, *rev*, *nef*) (Figure 4) [159, 160].

4.2 THE ENVELOPE GLYCOPROTEINS

The envelope glycoprotein (Env) complex, composed of gp120 and gp41, is the only virus-encoded determinant present on the HIV-1 surface. As such, it is the sole target for

neutralizing antibodies and of great significance for HIV-1 vaccine design. Because HIV-1 causes a chronic infection and currently circulating strains are the result of years of virus evolution in millions of hosts, Env has evolved extensive mechanisms to evade Ab responses, including glycan- and conformational shielding. In addition, the native Env complex is metastable and dissociates easily, making it an extremely challenging vaccine antigen [161].

4.2.1 Biosynthesis of Env

In the infected cell, the Env glycoprotein spike is synthesized as a single 90-kDa-polypeptide precursor consisting of ~847 to 875 amino acids translated from a 4.3kb spliced mRNA encoding the *env* gene [162]. In the rough endoplasmic reticulum (rER), this precursor polypeptide undergoes co-translational modification through the addition of 25-30 N-linked glycans to the protein backbone, resulting into the formation of a glycoprotein with a molecular weight of 160 kDa, termed gp160. The gp160 precursor protein assembles into trimers [163-167], which are transported to the Golgi apparatus where they are cleaved by host cell-encoded furin protease, generating the mature envelope surface glycoprotein, gp120, and the transmembrane glycoprotein, gp41. Trafficking through the Golgi apparatus allows processing of the glycans into a mixture of high-mannose and complex type N-linked sugars [168]. Following proteolytic cleavage, the gp120 and gp41 subunits remain associated through weak, non-covalent interactions, expressed as heterodimeric trimer complexes on the surface of infected cells [169, 170]. These gp120-gp41 complexes are incorporated into budding virus particles, where they are displayed on its surface as a functional trimeric spike. There is, however, the possibility that non-functional forms of the envelope spike, such as monomers or dimers could also be present on the virus surface as an outcome of either the dissociation of non-covalent gp120-gp41 complexes or inefficient processing of the spike during biosynthesis.

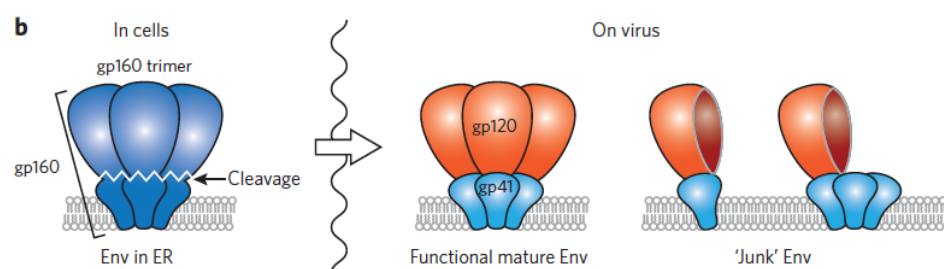


Figure 5: Schematic of Env glycoprotein variants after proteolytic cleavage of gp160 precursor polypeptide (adapted from ref [171])

Furthermore, the labile nature of the gp120-gp41 complex leads to shedding of gp120 from virus particles, generating immunogenic, non-native Env conformations or ‘decoys’ while leaving previously buried determinants in the gp41 ectodomain exposed on the viral surface (Figure 5) [172-174]. It was shown that only low levels of functional spikes (~10 spikes/virion) remain embedded in the virus particle making the spike density unusually low compared to other viruses [175, 176]. Nevertheless, this is clearly

sufficient for the infectivity of the virus given its successful spread in the human population.

4.2.2 Structure of Env spike

During HIV-1 infection, the surface-exposed gp120 subunit of the functional spike mediates attachment of the virus to CD4-expressing target cells, whereas gp41 is responsible for the fusion of the virus and target cell membranes.

The first glimpse of the HIV-1 Env structure, presented as a crystal structure of a monomeric gp120 core in complex with soluble CD4 and the 17b Fab (directed against the co-receptor binding site or gp120), was reported by Kwong et al and Wyatt et al in 1998 [177, 178]. These early structures uncovered the overall architecture of the CD4-bound form Env. Subsequent to this, several structures of gp120 cores in complex with different antibodies and/or ligands [179-189], as well as one structure of gp120 in its unliganded form [190]. The overall organization of gp120 in all these structures is highly similar, consisting of an inner domain (ID) that faces the trimer axis and gp41, an outer domain (OD) that is mostly exposed on the surface of the trimer and is heavily glycosylated, and a four-stranded bridging or β -sheet that forms the link between ID and OD. The primary receptor-binding site, the CD4bs, is located at the interface between ID, OD and bridging sheet (Figure 6).

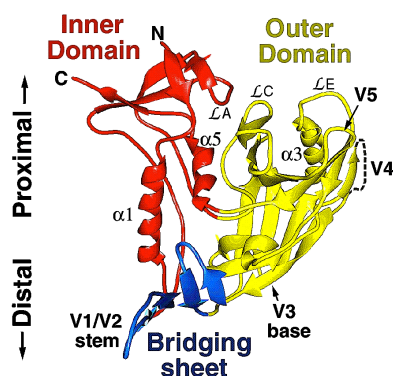


Figure 6: The ribbon diagram of the gp120 core from the perspective of the CD4 glycoprotein receptor. Inner domain shown in red and thought to interact with the gp41. Outer domain, which is variable and heavily glycosylated shown in yellow color. The linkage between outer and inner domain is bridging sheet shown in blue (adapted from ref [177])

Despite the success in resolving several structures of the gp120 core and gp41 subunits, stabilization and crystallization of intact Env spike in its native, pre-fusion state remained challenging. For a long time, the non-covalently linked metastable conformation of the Env complex, insolubility of membrane-bound proteins, extensive and variable glycosylation, and the presence of dynamic domains, hindered efforts to determine the molecular structure of the native, trimeric spike complex. However, using the SOSIP design (reviewed in [119]) in the context of the BG505 strain, a clade A transmitter/founder HIV-1 virus isolated from a Kenyan infant [191], well-ordered soluble trimers were generated. The SOSIP modifications include i) an addition of a disulfide linkage (termed SOS), which allows retention of the natural cleavage site between the gp120 and gp41 subunits, ii) an I559P mutation (termed IP) that prevents gp41 from “springing”, and iii) the deletion of the membrane-proximal external region (MPER) of gp41 to improve homogeneity and solubility of trimeric protein. The resulting stable, homogeneous, and non-aggregating protein, referred to as

BG505 SOSIP.664 gp140, was found to be recognized by several bNAbs suggesting that its structure is close mimic of the native trimer since the epitopes of bNAbs are believed to be present only on natively folded trimers [192]. In 2014, Wilson and colleagues solved the crystal structure of the BG505 SOSIP.664 trimers in complex with a potent bNAb, PGT122 at a resolution of 4.7Å, revealing the prefusion state of gp41, the interaction between the gp120 and gp41 protomers, and the spatial arrangement of variable regions, V1/V2, V3, HR1 and HR2 domains [193]. In addition, Ward and colleagues determined the cryo-electron microscopy (cryo-EM) structure of the same Env variant in complex with a different bNAb, PGV04 at a 5.8Å resolution [194]. The two structures obtained with different methodologies (X-ray and cryo-EM) were similar to each other and consistent with a cryo-EM structure reported by Subramaniam and coworkers [195]. More recently, Guenaga et al. reported the high-resolution Env structure of subtype C primary isolate 16055 using trimers produced by a different strategy than SOSIP design, the native flexibly-linked (NFL) trimers, which are described in more detail in section 4.4.2 [196]. Understanding the structure of the native trimeric Env spike for a long time was considered to be a bottleneck for developing an effective HIV-1 vaccine, thus these were important improvements.

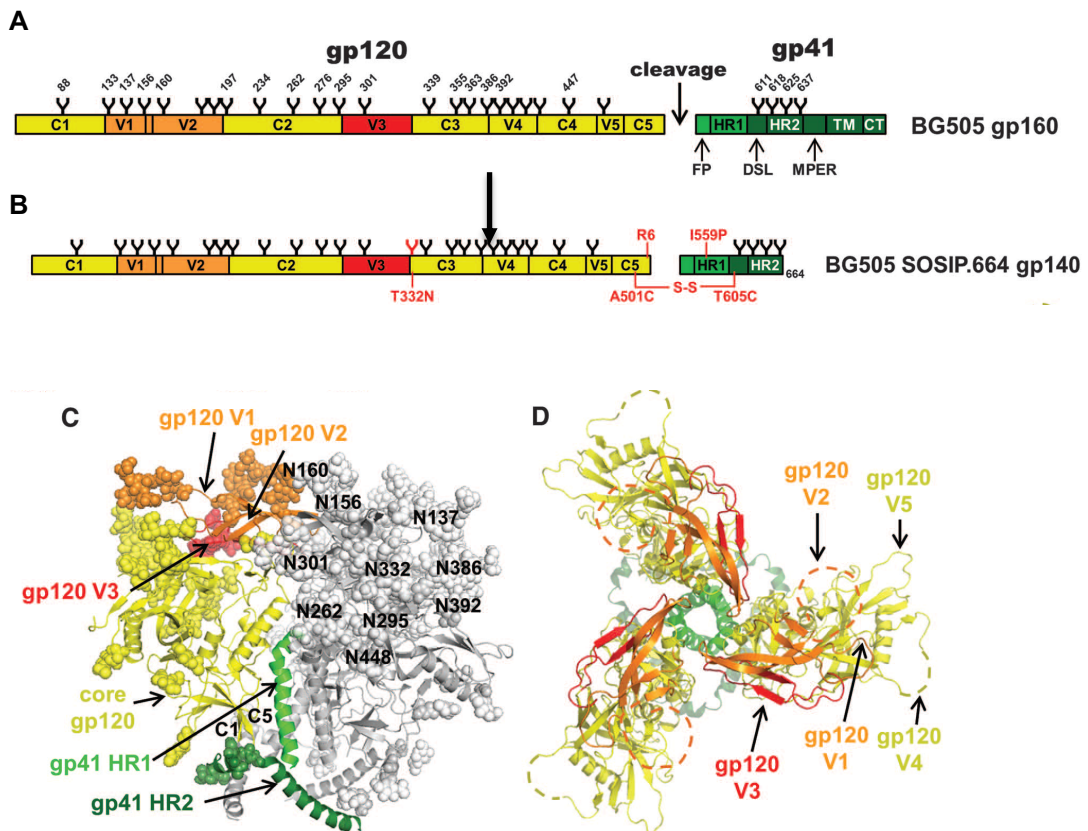


Figure 7: (A) Schematic of a full-length gp160 Env construct of HIV-1 isolate BG505. N-linked glycans are numbered on their respective Asn-residues. (B) The modifications in the native gp160 included SOSIP mutations and the addition of N332 glycan site (shown in red in B) to produce a stable gp140 trimer, referred as BG505 SOSIP.664. (C) Three-dimensional structure (side view) of the BG505 SOSIP trimer. The gp120 core is shown in yellow, variable V1/V2 and V3 regions are shown in orange and red,

respectively. The gp41 elements are highlighted in green and glycans are represented as grey spheres on one of the three protomers of Env. (D) Top view (down to the trimer axis) of Env. Highly variable loops V1 to V5 in gp120 mapped to the periphery of the trimer are shown where dashed lines for V2 and V5 indicate the absence or ambiguous electron density (adapted from [193]).

4.3 ANTIBODY RESPONSES TO HIV-1 ENV IN INFECTION

HIV-1 Env is presented to the humoral immune system in several different configurations in chronically infected humans, including the infectious native form on the surface of virions, shed monomeric gp120, gp120-depleted gp41 stumps on virions and various forms of unprocessed forms of gp160 released from the cytoplasm of infected cells [197, 198]. Following HIV-1 infection, anti-Env Abs are detected within ~1 week of detectable viremia [199], primarily against the gp41 ectodomain of Env. Following production of circulating anti-gp41 Abs, ~2 weeks later, anti-gp120 Abs that target variable region 3 (V3) appear. However, these binding Abs are not neutralizing and do not affect viremia detectably [199]. Consequently, they exert little or no selective pressure on the Env [200]. The first neutralizing Abs (NAbs) against the infecting strain (autologous virus) are detectable several months post-infection [201-204]. Although autologous NAbs are not effective against the divergent (heterologous) viruses, they can exert significant pressure on the circulating virus [205], which promotes neutralization escape by selecting for variants that differ by single amino acid substitutions, insertions, deletions or by the position of N-linked glycans, shaping the evolution of the virus as the infection progresses [206-208].

4.3.1 Broadly neutralizing antibodies (bNAbs)

The development of cross-reactive NAbs, also called broadly neutralizing Abs (bNAbs), capable of neutralizing a broad range of heterologous HIV-1 variants generally takes years of chronic HIV-1 infection [209-212]. Approximately 20-30% of HIV-1 infected subjects were shown to develop moderate to highly cross-reactive Abs during the first three years of infection [204, 210, 212-216]. Among them, about 1% develop bNAbs with exceptionally potent and broad neutralizing activity, which can neutralize diverse HIV-1 strains of multiple subtypes. [211, 217, 218]. Recently it was shown that infected infants could produce bNAbs 20 months after infection, indicating that genesis of bNAb responses is not restricted to the adults [219]. Mapping of broadly neutralizing sera revealed a limited set of target epitopes [220-222]. It is generally believed that if a vaccine could stimulate bNAbs it would be likely to confer protection against HIV-1 as highly potent bNAbs isolated from infected individuals [223, 224] are capable of preventing infection in macaques or suppressing viremia in a mouse model of HIV-1 infection upon passive transfer [225-229].

The isolation and understanding of the complex evolution/maturation pathway of these bNAbs, along with their molecular interactions with the HIV-1 Env at the atomic level, have been possible due to recent technical advancement in single-cell Ab cloning, high-throughput sequencing, structural biology, and protein engineering [230]. Several of the Ab-Ag complex structures resolved to date have revealed important information about the conserved surfaces that are present and accessible on the native Env trimer. The information about the ‘sites of vulnerability’ or ‘supersites’ such as the CD4-binding site (CD4bs), the trimer quaternary and glycan/glycopeptide epitopes on gp120, and the membrane-proximal external region on gp41 have implications for immunogen design [231-234].

4.4 ENV TRIMER AS A TEMPLATE FOR HIV-1 VACCINE DESIGN

4.4.1 Early generation Env trimers

An effective prophylactic HIV-1 vaccine must provide protection against the enormous variety of circulating HIV-1 strains, and if not sterilizing, must curb the infection before the formation of reservoirs. Vaccines developed against HIV-1 to date have shown little or no success in more than 200 vaccine trials conducted worldwide [235]. In many of these trials, the vaccine antigen was derived from the HIV-1 envelope monomer glycoproteins gp120, and/or gp41, however these failed to induce potent and broad neutralizing Ab responses against diverse HIV-1 strains [236, 237]. After the failure of gp120 monomers as an effective vaccine, the focus of immunogen design shifted towards the development of soluble trimeric mimics of the Env spike. The generation of soluble Env trimers requires the construct containing both gp120 and gp41 subunits. The secreted/soluble proteins are easy to produce and purify with a good yield compared to their membrane-anchored counterparts. Thus, one of the early approaches was to introduce a stop codon at the C-terminus of the gp41 ectodomain prior to the transmembrane domain to allow the truncation of the gp41 subunit, resulting in soluble gp140 protein that contain the gp120 subunit and the ectodomain of gp41. However, this genetic truncation adversely affected the stability of the already metastable trimer leading to misfolding and dissociation of gp140 into the constituent gp120 monomer and trimeric gp41ectodomain (reviewed in [238]). This issue was later addressed by either inactivating via point mutations or deleting the cleavage site of the gp160 precursor. As a result, the association between the gp120 and gp41 subunits was improved, although the resultant secreted Env trimers were highly heterogeneous and consisted of a mixture of multiple gp140 conformations. The next step in Env engineering involved the addition of exogenous trimerization motifs, for example, foldon or GCN4, to the C-terminal of gp41 ectodomain, to obtain more stable and homogenous Env trimers. This modification improved the recovery of the Env fraction and enabled the production of Env immunogen in sufficient amounts to perform immunogenicity studies in animal models [239, 240].

One of this type of soluble Env trimers, gp140 foldon (gp140-F), derived from the clade B HIV-1 isolate YU2, was used in **papers I and II** of this PhD thesis. Soluble YU2 gp140-F immunogen has been extensively characterized for its antigenic and

immunogenic properties in small animals [241-244] as well as in non-human primates (NHPs), providing a wealth of important immunological information about the humoral immune responses in the context of Env vaccination [32, 122, 124, 245-248]. Although this first generation gp140-F trimer was able to induce robust B cell responses in NHPs [248], they failed to elicit the antibodies that were capable of neutralizing a broad panel of circulating HIV-1 strains (also called as Tier 2 viruses) ([32], **paper I**). In recent years, following the introduction of new analytical methods to study the molecular conformations and biophysical properties of Env trimers, it has become clear that gp140-F trimers are not faithful mimetics of the native Env spike and that they have an open conformation exposing several non-broad neutralizing epitopes [124, 249-254]. While the foldon trimers were designed empirically in the absence of a high-resolution crystal structure of the native Env spike, they represented a state-of-the art Env design at the time, displaying high homogeneity by standard biochemical gel and chromatography analysis. The use of these trimers in immunogenicity studies provided important initial information about the elicited B cell responses and also helped establishing several analytical tools and reagents that are broadly used in the field today [32, 122-125, 248, 255, 256].

4.4.2 Next generation native-like Env trimers

As described in section 4.2.2, the two key substitutions, 1) the introduction of cysteine residues in each of the gp120 and gp41 subunits to covalently link gp120 to gp41 (501C-605C; referred to as SOS) and, 2) the insertion of a helix-disrupting point mutation, I559P, in heptad repeat 1 to prevent post-fusion conformational changes, formed the basis for one of the new generation Env trimers. Additional modification by truncating the MPER of the gp41 resulted in a soluble, cleaved, well-ordered trimer called BG505.SOSIP.664. However, the SOSIP modifications when applied to Env from different HIV-1 strains resulted in a mixture of native-like and non-native like forms (reviewed in [119]), indicating that Env from the BG505 had an intrinsic ability to fold into native-like trimers, unlike other strains.

In the meantime, an alternative approach developed by Wyatt and co-workers involved the replacement of the furin cleavage motif REKR with a flexible glycine-serine based (G4S) linker at the interface of the gp120 and gp41 subunits to covalently link these two subunits together. This design also included the introduction of the I559P point mutation, resulting in soluble, cleavage-independent homogenous mimics of the native Env trimer, the NFL trimers [249]. The NFL design when applied to the BG505 Env formed well-ordered trimers resulting in homogenous trimer preparations, while Env from strains such as JRFL (clade B) and 16055 (clade C) produced a mixture of well-ordered and non-well-ordered trimers, aggregates, dimers and monomers. This demonstrated that BG505 Env is an exception in its ability to form a high percentage of native-like trimers by either the SOSIP or NFL modifications. To obtain well-ordered trimers from a non-homogenous complex mixture, positive selection using trimer-specific bNAbs such as PGT145 or VRC26, or negative selection to remove disordered trimers by a non-bNAb such as the CD4bs-directed MAb, F105 were used, however, the percentage recovery of

well-ordered trimers was low [249, 250]. To increase the yield of well-ordered trimers more sophisticated approaches, such as structure-guided enhancements, were applied to the initial NFL trimer design. Using BG505 structural information, 8 or 15 amino acid residues near the gp120-gp41 interface that are involved in trimer stability, referred as trimer-derived (TD), were identified and transferred to the 16055 or JRFL Envs, respectively. The resultant trimer is termed as 16055 NFL TD or JRFL NFL TD. Further optimizations were made, including introduction of a cysteine residue pair (I201C-A433C) to prevent CD4-induced conformational changes after gp120:CD4 interaction (trimers referred as NFL TD CC). Furthermore, glycine substitutions in the gp41 region were introduced to obstruct conformational transitions to post-fusion, helix-dominated conformations (trimers referred as NFL TD CC (T569G)). These modifications increased the overall stability and homogeneity of the trimers and resulted in higher yields [196, 257]. Recently, Wyatt, Wilson and collaborators solved the crystal structure of the NFL BG505 Env trimers, in complex with two potent bNAbs (PGV19 and PGT122), at 3.39 Å resolution. These studies demonstrated that the NFL and SOSIP designs are highly similar at the structural level and in their glycosylation profiles. In addition, the flexible linker that connect gp120 and gp41 units in NFL trimers was found to maintain native-like prefusion compact structure without distorting known epitopes on Env or affecting the inter-domain flexibility, similar to SOSIP constructs [258].

The highly stable, well-ordered NFL or SOSIP trimers enable the use of additional modifications such as targeted deletion of N-linked glycans proximal to the conserved CD4bs to increase exposure of this principle neutralizing epitope for BCR access and B cell activation, without altering other determinants of the well-ordered trimers [259, 260]. Such glycan-deleted trimer variants on the 16055 Env background were shown to prime the neutralizing Ab responses, efficiently in rabbit immunogenicity experiments [259]. In addition, soluble native-like Env trimers or glycan deleted variants can also be arrayed on particles to generate virus-like particles (VLPs) [261] or on synthetic liposomal nanoparticles to improve immune response by activating more B cells. For example, Env trimer-conjugated liposomes were superior in activating Env-specific murine B cells *ex vivo* and more efficient in inducing GC B cells *in vivo*, compared to soluble Env trimers [262]. In NHPs, trimer-arrayed liposomes were also shown to induce superior GC responses as well as to generate higher neutralizing Ab titers against the autologous tier 2 16055 virus compared to the same quantity of strain-matching soluble trimers in adjuvant [121]. So far, the native-like Env trimers were shown to induce strong neutralizing Ab titers against the autologous tier 2 viruses in the relevant animal models; however, potent heterologous neutralization against a broad spectrum of circulating HIV-1 strains is yet to be achieved.

4.5 ANALYSIS OF ENV TRIMER VACCINE-INDUCED B CELL RESPONSES

The success of a vaccine relies on the ability to induce robust, long-term protective adaptive immune responses. These responses comprise memory B cells, T cells and long-lasting Ab-producing cells, which differ in terms of their quantity, quality, homing properties, and persistence over time depending on the vaccine formulation (antigen,

adjuvant, delivery system). An in-depth understanding of B cells responses, the main contributor of protective adaptive immunity, is essential for the development of effective vaccines against a specific pathogen such as HIV-1. In recent years, a rapid development and expansion of analytical tools have enabled higher resolution interrogation of pathogen-specific B cell responses after infection or vaccination. Nevertheless, the comprehensive analysis of vaccine-elicited B cell responses in humans is difficult due to the inaccessibility of tissues including bone marrow, gut, and spleen where key B lymphocytes population such as LLPCs resides. However, this is possible in NHPs that are used frequently in biomedical research to model human immunology owing their similarity to humans both at the genetic and cellular level. Immunogenicity studies using early generation Env trimer in NHPs have enabled us to dissect the B cell responses at the cellular, functional, and genetic level in different compartments, providing highly detailed information about induced B cell responses ([32, 121, 248], paper I, III, IV).

4.5.1 At the cell subset level

Our group previously showed that inoculation of Env administered in adjuvant elicits robust peripheral Env-specific IgG plasma cell and memory cell responses in rhesus macaques [248]. Similar response kinetics were observed for another recombinant glycoprotein, the influenza virus hemagglutinin (HA) [245]. In more recent immunogenicity trials we found that, while the kinetics of induction and contraction followed the same pattern as observed with the first generation trimers, the overall magnitude of the response was lower upon vaccination with new generation native-like Env trimers ([121] and **paper IV**). The qualitative difference between B cell responses in these two studies was likely due to the Env trimers used for vaccination, wherein new generation native-like Env trimers are more tightly packed and do not expose non-neutralizing or non-broadly neutralizing determinants, resulting in a lower overall magnitude of induced Ab response [121, 248]. The majority of the LLPCs are bone marrow (BM) resident, where they can persist for decades in the absence of antigen [101] and are responsible for the production and the maintenance of high-affinity antibodies and therefore provide immediate protection upon antigen re-encounter. ELISPOT analyses of B cell responses in the BM have shown the different kinetics of Env-specific BM plasma cells ([121, 248] and **paper IV**).

4.5.2 At the single cell level

While analyses of different subsets of B cells have provided important information about the kinetics and magnitude of Env-trimer specific Ab responses, more details about the vaccine-induced B cell responses can be obtained by isolating antigen-specific B cells at the single cell level. Analysis of Ab responses at the clonal level using single cell RT-PCR expression cloning enables the examination of the underlying genetic diversity of the response and, through the isolation of MAb, gives opportunities to study functional and structural properties of individual antibodies specific for distinct epitopes on the antigen (reviewed in [54]). Several years ago, our group established methods to isolate

Env trimer-specific MAbs from single memory B cell obtained from immunized rhesus macaques. In this thesis, I build on these studies by combining MAb isolation with NGS immune repertoire analysis.

5 AIMS

The specific aims of the individual papers were:

- Paper I:** To determine the antibody VH gene usage in the overall B cell repertoire and in response to HIV-1 Env vaccine in rhesus macaques.
- Paper II:** To investigate antibody responses against HIV-1 Env by isolation and evaluation of genetic and functional properties of a panel of monoclonal antibodies (MAbs) from immunized rhesus macaques.
- Paper III:** To develop a program for *de novo* identification of germline V genes to construct individualized V gene databases and to improve the current rhesus macaque immunoglobulin germline gene database.
- Paper IV:** To study the diversity, distribution and maturation of vaccine-induced Env-specific B cell lineages in the multiple immune compartments of immunized macaques.

6 SUMMARY OF PAPERS

A hallmark of the humoral immune system is the diversity of our Ab repertoires that can recognize virtually any antigen. For decades, Ab responses to infections or vaccines were monitored primarily by serological assays that assess Ab specificities, but which do not reveal the underlying dynamics in clonal populations of B cells, or the specific Ig gene rearrangements coding for the antibodies. In recent years, the advent of MAb isolation methods and high-throughput DNA sequencing has enabled the interrogation of Ab repertoires to a higher level of detail, providing greater insight into the molecular details of genes encoding thousands to millions of antibodies. In this thesis, I developed and used methods to interrogate B cell responses at both genetic and functional level in the context of HIV-1 Env vaccination using in a highly relevant animal model for understanding human B cell biology, rhesus macaques. We developed a novel computational tool called IgDiscover to identify germline Ig alleles (both known and previously undefined) present in a given subject, enabling the construction of personalized/individualized reference Ig gene databases. The use of IgDiscover to produce individualized antibody germline gene databases from outbred populations is revealing considerable diversity between subjects, providing new knowledge about the immunoglobulin allelic variation in a given species. Importantly, the construction of individualized reference databases allows accurate antibody gene assignment, expression profiling, SHM estimation and antibody lineage tracing in a given subject. Here, I used these methods to comprehensively characterize the genetic diversity, clonal distribution/dynamics and maturation of B cell responses induced by HIV-1 Env trimer vaccines. In this section, I briefly present the results of **papers I-IV**.

6.1 PAPER I:

Previous work from our group has showed that the germline Ig gene segments encoding the Ab heavy and light chains of macaque Abs are highly similar to their human counterparts and cluster according to gene family, rather than species, suggesting a high degree of conservation in Ig loci [32]. However, detailed knowledge about Ig genetics in healthy and vaccinated rhesus macaques was lacking. Thus, in **paper I**, to gain insight into B cell responses at the genetic level, we studied B cell repertoires by using two approaches, 454-pyrosequencing and single-cell RT-PCR of bulk and sorted memory B cells, respectively. In particular, we investigated the contribution of individual V gene segments in total IgG-switched and Ag-specific B cell repertoires from Env-immunized rhesus macaques. The mapping of resultant V(D)J sequences to the reference database revealed that the overall expression of Ig gene segments in the major VH families (VH1, 3, and 4) was unequal with few expressed abundantly, whereas others were utilized at rare or undetectable frequencies. A similar pattern of gene usage was shown in human following 454-pyrosequencing of the human Ab repertoire [263]. Importantly, the expression of gene segments detected by both Sanger sequencing and 454-pyrosequencing methods gave remarkably concordant results despite performing the PCR under different condition, indicating that there was no major bias in amplifying certain gene segments by the PCR protocols used prior sequencing by either of the

methods. In addition, we also investigated the V gene usage in Ag-specific memory B cells elicited in response to an experimental protein-based Env vaccine. Specifically, we examined >1000 sequences generated from single sorted Env-specific and IgG-switched total memory B cells from macaques immunized with well-characterized recombinant HIV-1 Env trimers. The memory B cell response to Env was highly polyclonal, consisting of diverse Ab-encoding germline V genes, similar to the V gene usage in the IgG-switched total memory B cell repertoire. Clonal grouping of a total of 606 Env-specific sequences estimated that 502 unique Ab clonotypes were present in the Env-specific memory B cell response, indicating broad diversification of Ag-specific memory B cell repertoires upon Env vaccination. The highly polyclonal B cell response we observed in our study was similar to that observed following immunization with protein-based tetanus toxoid in humans, which consisted of ~100 unique Ab clonotype, [264]. In contrast to the diverse vaccine-induced B cell response to HIV-1 Env vaccination observed in this study, chronic HIV-1 infection was shown to elicit a more restricted memory B cell responses consisting of only ~50 clonotypes with a biased usage of IGHV1 family gene segments [265], suggesting that skewing of the VH gene usage is a consequence of persistent Ag exposure.

6.2 PAPER II:

In **paper II**, I extended the genetic analyses of Env-specific memory B cells, studied in **paper I**, to the functional characterization by isolating a panel of vaccine-induced MAbs. Specifically, to elucidate which Ab sub-specificities were archived in memory B cells elicited by HIV-1 Env trimer vaccination, we cloned and expressed 52 MAbs from single sorted Env-specific B cells from two NHP donors. Using differential ELISA with various Env ligands I could define distinct sub-determinants of Env the MAbs recognized, including the primary receptor binding site (CD4bs), the variable region-1/2 and 3 (V1V2 and V3), gp41, and the foldon trimerization motif. The germline V, D, and J gene segment assignment to Ab heavy and light chains revealed broad gene usage among the different Env-specific MAbs, indicating no bias towards the usage of a given gene segment against any of the epitopes. The level of somatic hypermutation in V genes of most of the MAbs was low. The relatively low levels of SHM we observed in vaccine-induced MAbs isolated in this study is in contrast to that observed for MAbs isolated from HIV-1 infected patients, which are highly mutated, likely as a consequence of persistent Ag exposure due to replicating and mutating virus during chronic HIV-1 infection. Vaccine-induced MAbs are generally less mutated than infection-induced MAbs, as shown in **paper IV**, in our previous study [32], and other studies of vaccination [264]. Next, all 52 MAbs were evaluated for their virus neutralizing capacity in the *in vitro* neutralization assay using a panel of commonly used HIV-1 Env pseudoviruses. Among all the MAbs, 25% targeting two sub-determinants of Env, CD4bs and V3 region were neutralizing and showed distinct neutralization signatures depending on their sub-specificity, whereas no neutralizing activity was detected for the gp41-, foldon-, or V1V2-specific MAbs. The neutralizing activity of the CD4bs- and V3- directed MAbs was in agreement with the unfractionated plasma samples, suggesting that they are the major specificities archived in the memory B cell pool. Interestingly, while neutralizing

activity of the plasma against some viruses was undetected, purified MAbs from the same subject showed considerable neutralization against those viruses, indicating the importance of MAb isolation to obtain higher resolution analysis of induced B cell responses. In summary, isolation and detail characterization of vaccine-induced MAbs in this study have provided improved knowledge about the limitations of existing immunogens to induce broad protective neutralizing Abs against hypervariable pathogen such as HIV-1. We anticipate that in the near future, the systematic analysis as we showed here and isolation of additional MAbs elicited by different candidate immunogens and immunization protocols will generate informative results that will guide effective vaccine discovery efforts against HIV-1.

6.3 PAPER III.

Comprehensive knowledge of immunoglobulin germline genes is essential to better understand the genetic basis of B cell responses. In **paper I** and **paper II**, to study the genetic properties of primate B cell responses I used our previously annotated and published reference germline database [32], which was derived from one female Indian rhesus macaque genome. While those studies provided a great deal of baseline information about the genetic diversity of Ig genes in Ag-specific B cell responses, the total expressed Ab repertoire in rhesus macaques suggested that the existence of additional, hitherto undescribed V gene germline sequences, as well as considerable variation between individual animals. Thus, we realized that the use of non-human primate species for immunoglobulin gene repertoire analyses would benefit greatly from the availability of more complete germline V gene databases. To achieve this, we set out to devise a method to elucidate individual germline gene repertoires of rhesus macaques to facilitate studies of Ab genetics and affinity maturation in outbred populations. In **paper III**, we developed a novel approach combining HTS of expressed IgM repertoire with bioinformatic analysis of IgM sequences to identify germline V genes in a given subject. The resultant program, called IgDiscover, allows the identification of both known and novel germline sequences, enabling the *de novo* construction of individualized reference V gene database. Other computational approaches to identify the novel Ig alleles or detect the Ig polymorphism were reported [56, 266]. However, until the development of IgDiscover (**paper III**), a reliable, rapid and automated method to construct subject-specific germline V gene database was not available, especially for species that lack a complete reference genome or have incomplete reference sequences in regions of high genomic complexity such as the Ig locus. The ability of IgDiscover to produce individualized databases with little or no prior genomic information allows both inter- and intra-species comparison has broad implications for the field of immunology. For example, production of individualized reference databases from Indian- and Chinese-origin macaques revealed a high degree of allelic diversity in their V genes. The level of Ig allelic diversity we observed in macaques was confirmed recently by a study comparing Ig alleles obtained from genomic sequencing of the Ig loci from nine macaques [33]. Importantly, the availability of definitive reference for each subject improves gene assignment and thus, accurate immune profiling and SHM estimations, which will accelerate our understanding of V gene usage and selection pressure

experienced by particular gene segments during immune reactions.

6.4 PAPER IV.

In **paper IV**, I dissect the B cell responses induced by the NFL Env trimers in multiple immune compartments and at multiple time points to understand the diversity, the distribution/dynamics and the maturation of hundreds of Env-specific lineages. First, using IgDiscover we generated an individualized heavy chain germline V gene database for the animal selected for detailed investigation. Individualized V gene database is necessary for correct genes assignments of Ab sequences, SHM estimations, Ab lineage assignments and tracing, as described in section 3.3. We next developed two extension modules to the IgDiscover tool, *Clonotypes* and *Clonoquery* to interrogate total and Ag-specific B cell repertoires at the quantitative and qualitative levels, respectively. We detected similar V and J gene usage patterns in pre- and post-immunization samples, suggesting that Env vaccine-induced response comprise a minor fraction of total memory B cell repertoire and consistent with previous knowledge obtained from B cell ELISpot analyses [248, 256]. To investigate how individual antigen-specific B cell lineages expand and disseminate in different immune compartments, we combined single B cell sorting, MAb isolation and high-throughput IgG repertoire sequencing of post-immunization samples from blood, bone marrow, spleen, draining lymph node, and gut (ileum). We sorted over 1000 Env-specific single memory B cells, which comprises over 200 unique Ab clonotypes based on identical V, J gene usage and identical HCDR3 region, suggesting highly polyclonal Env-specific response. The *Clonoquery* based high throughput clonal lineage analysis revealed broad dissemination of the vaccine-induced B cell clonal lineages in blood, draining lymph node, spleen and bone marrow. We demonstrated that the level of SHM at the intra-clonal level across different time points and sites was different. We found that affinity-matured Ab variants generated in response to later boosts replaced or outnumber previous variants from the same lineage in both blood and bone marrow. In addition, we detect the large number of variants for given B cell clonal lineages in draining lymph node and spleen, which are the sites of ongoing GC reaction as we found mixture of Env-specific memory and plasma cells by B cell ELISpot analysis. In contrast, in blood and bone marrow, which comprise the final output from the GC reaction, using ELISpot assay we detect memory B cells and plasma cells, respectively. Furthermore, we analysed the gut (ileum) sample for the presence of Env-specific B cells by both B cell clonal lineage tracing and ELISpot analysis. Both methods gave highly comparable results with minuscule or no detection of Env-specific B cells in the gut (ileum), suggesting that the vaccine response disseminates to this compartment poorly. The poor distribution of Ag-specific responses we observed in gut is consistent with the minimum clonal overlap that we observed between the IgG repertoires from gut and blood or bone marrow and similar to a recent study of total B cell repertoires in humans demonstrating that the gut B cell repertoire is qualitatively different from the blood B cell repertoire [267].

7 CONCLUDING REMARKS

The aim of this thesis was to analyse vaccine-induced B cell responses induced by HIV-1 Env glycoprotein trimer immunization to gain improved understanding of basic B cell biology, as well as to guide the design of future candidate vaccines against HIV-1. At the beginning of this PhD thesis, very little was known about the Abs elicited by Env vaccination in a relevant animal model such as rhesus macaques with no molecular understanding of the induced Ab repertoire that these Abs are part of. This was mainly due to insufficient information about antibody germline genes in rhesus macaques, limiting our knowledge about the genetic makeup of B cell repertoires in this important animal model.

The present thesis defines the germline repertoires in rhesus macaques as well as comprehensively characterizes the vaccine-induced Ab repertoires in multiple immune compartments at high levels of detail. **Paper I** in this thesis interrogates the total and Env-vaccine induced B cell repertoires to comprehend the underlying genetic basis of B cell responses. The diverse Env-specific genetic response that we observed was comparable to total Ab repertoire, providing a first baseline about the genetic composition of vaccine-induced B cell responses against the complex protein antigen. The dissection of polyclonal sera at the level of monoclonal antibodies in **paper II** allows characterization and comparison of the fine Ab specificities that are archived in the memory B cell pool following Env immunization. The genetic diversity of MAbs targeting the same epitope helps to improve our basic understanding of B cell responses. Understanding the functional composition by mapping MAbs to their cognate epitopes has important implications in designing improved immunogens. Furthermore, the IgDiscover tool developed as a part of this thesis in **paper III** revealed important information about allelic diversity in Ig genes in rhesus macaques. The striking Ig allelic diversity we observed highlights the necessity of individualized Ab germline database to avoid erroneous immunological conclusions drawn from Ab repertoire data and to improve the quality of B cell repertoire studies. Importantly, the production of individualized germline database opens new avenues to understand Ig allelic diversity in human and other species. The resulting information will facilitate the assessment of previously unknown associations between immunological profiles, disease susceptibility and/or immunization efficiency and efficacy. Finally, dissemination and maturation of hundreds of vaccine-induced Ab lineages across multiple immune compartments that I studied in **paper IV** provide a first anatomic atlas of antigen-specific B cell clonal lineages elicited in response to Env vaccination.

In conclusion, the highly detailed information about vaccine-induced B cell responses against real world antigens such as HIV-1 Env used in this thesis not only updates our knowledge of B cell biology in the context of vaccination but also guides to develop improved immunization regimens.

8 ACKNOWLEDGEMENTS

During my PhD time, I had the pleasure to work and interact with many great people who helped me grow and broaden my scientific knowledge and skills. I am truly grateful to those who have contributed to this thesis in one way or another:

First and foremost, my supervisor, **Gunilla Karlsson Hedestam**, you are the best teacher and mentor in my entire academic training. You are truly inspiring and I will forever be grateful for your contribution to my scientific as well as personal development. My profound thanks for your trust, freedom, constant support and encouragement during these fantastic 5^{1/2} years of my life.

My co-supervisor, **Mats Persson** for your encouragement and help in getting the important WABI support. Thank you for providing the information on various workshops and programs related to antibody repertoires, high-throughput sequencing and data analysis throughout my PhD.

My co-supervisor, **Christopher Sundling** for your great supervision during the early days of my PhD. It is true pleasure working with you and I always enjoy discussing science with you. You are very nice and easy-going person always there to help when needed. Thank you for your scientific inputs, advices, social times and setting the bar high! I hope to work/collaborate with you more in the future.

Our main collaborator, **Richard Wyatt** for your immense knowledge on Env and for crucial collaboration. It is real treat and enriching experience to listen science from you. Thank you for excellent scientific inputs and advices on our projects and for enlightening us about Env every time you visited us.

My Mentor, **Carl Johan Sundberg** for your wisdom and guidance.

Members of the GKH laboratory

I am truly thankful for having had the chance to work with such great team members who are creative, collaborative, fun, professional and supportive. Undoubtedly, we make one of the best teams in KI. I would like to extend my warm thanks to - **Paola**, for being in the PhD journey together, your hard work and expertise in cellular techniques was indeed encouraging and helpful, thank you for your encouragement during my lonely days of 450bp analysis, for giving me good advices in the difficult times. Also, you are good at maintaining the right 'work-life' balance, which I'm trying to imbibe. **Martin**, for all the discussions and brainstorming sessions to develop IgDiscover, for early MiSeq runs together, your help in preparing me for the big talks, pedagogic lab meetings and the sarcasm. Your dedication at work is motivating. Keep your cool! **Pradeepa**, for your cheerful personality, patience while working with me, for working independently and your significant contribution to paper IV in this thesis. You will do great in the rest of your PhD journey. Congratulations and happy parenting ahead! **Monika**, we started in

the group on the same day and you are definitely one of the best colleagues I've worked with so far. I admire your calmness. You are a great team player and always know when to lend a hand. Thank you for being thorough, for teaching/helping with FACS experiments, for relaxing ping-pong matches and discussions on life and science during the games. (Also, thanks to **Csaba** for your perspective on life and sharing the story about Rabindranath Tagore). **Martina**, for being a kind and very helpful person inside and outside the lab. You are organized, easy going and witty – it was pleasant to have you around! Thank you for all the fun at meetings/conferences, memes and giggles. **Néstor**, for your energy at work, your willingness to learn new things and for all your help on the long sampling days, not to forget, the great photoshop skills. Keep up the good work! **Elina**, for being a good listener and a level-headed person. It is fun to hang out with you during lunch, coffee breaks etc. Thank you for sharing the great news and sorry for missing on it, hugs! **Sharesta**, for great work and the talks. You are a bundle of energy, always on the go – be it tough Viking race, bouldering or gym sessions. Thank you for being so kind to give me neck massages during hours of data analysis. **Sanjana**, for your help with R scripts and for kindly offering me food from your lunch box. Newest member of the lab, **Uta**, for your willingness to help, good discussions and questions in the lab meetings. **Marco**, for being a nice person and always offering the cookies to me and to everybody.

Gabriel, for intelligent questions/inputs in the lab meetings, your relaxed personality and light humour creates smooth atmosphere at work, and thanks for a great time in Banff! **Marjon**, for early supervision in the lab, giving good advices on PhD studies and for occasional social times at Mosebacke. **Pia**, for renting your apartment to me during early days in Stockholm – I had a pleasant stay. Thanks to you and **Will** for taking me out for a dinner in NY. **Lotta**, for detailed protocols and for being a great pub buddy at MTC.

To students: **Nikos**, **Anna Klara**, **Komal**, **Lina**, and **Sigrun**, you all are so talented, and hard-working. Your contribution to this thesis is important and not to be forgotten. Thank you for giving one of the best experiences during my PhD journey and helping me to understand and appreciate Nilla's role as a supervisor.

To all former and present associated labs:

The McNerney group - **Gerry**, for your expertise in virology, for being a nice person and your wittiness. Thank you for giving me opportunities to teach and supervise the students through your research school programs. **Marc**, for your curiosity and tidiness. **Bastian**, for pedagogic presentations, for sending in the greetings on my birthday every year – Thank you! **Ben**, for the Brunnsviken times. **Leo**, for taking time to read the paper IV and being a part of the discussion. Previous and current members, **Kai**, **Roberta**, **Thomas**, **Lifeng**, **Ainhua**, for creating nice and friendly environment at work.

T cell team - **Jonathan**, for your help during my first paper submission, for providing good scientific inputs, explaining T cells during lunch for my book exam and for being a passionate individual in general. **Chris**, for an awesome time at winter school and your help with navigation in Japan. **Julian**, for being the human encyclopaedia. **Leona**, for being a fun company at occasional lab outings. **Moh**, for your chilled attitude. **Junjie**, for company on weekends at work and to a new member, **Akashdip**, thanks!

The Loré lab: **Karin Loré** for active collaboration and for your fantastic work on basic vaccine immunology. Previous members **Kerrie**, **Emily**, for friendly atmosphere at work and current members **Frank**, **Liz**, **Gustaf**, **Sebastian**, **Ang**, for collaboration, discussions and scientific inputs during the joint lab retreat.

Anna Smed Sörensen and group, for your DC expertise, **Saskia**, for being so organized and for the speedminton games, it was a pleasure sharing the office with you. **Faezzah**, for great work, nice presentations and for my first ski experience with you in Sälen.

To collaborators:

WABI - for providing an excellent bioinformatics support. I would especially like to thank **Marcel** for your awesome coding skills and for all the meetings, discussions, and brainstorming sessions that we have had to develop IgDiscover and other modules. Also, for providing elegant solutions to our countless requests in the issue list and for your patience with us. Your contribution to paper II and IV in this thesis is significant – Thank you!

Members of the Wyatt lab – For providing us with immunogens and reagents used in paper I, II and IV. **Yu Feng**, **Jidnyasa**, **Viktoriya**, **Shailesh**, **Shridhar**, **Salar**, for welcoming me and taking good care of me and **Javier**, for a brief pymol tutorial during my visit to TSRI. **Karen**, for your help in paper IV.

To collaborator at VRC, NIH: **John Mascola**, for your valuable scientific inputs on our studies and a special thanks to **Sijy O'Dell** for high quality neutralization assays and analyses of our NHP samples.

To the Astrid Fagareus lab at Karolinska Institutet: A special thanks to **Mats Spångberg**, **Bengt**, **Pia** and team for taking good care of our animals and for all long, intricate samplings with great expertise.

To **Xueling** and **Eric** at Aaron Diamond AIDS Research Center, NY- Thank you very much for hosting me and for my first introduction to MiSeq.

To **Jamie** and **Felix** at Simon Fraser University – For forming AIRR community and iReceptor initiative. Thank you so much for having me in your house and for an opportunity to present our research at SFU. Felix – your valuable feedback and inputs after the talk were helpful. And You are an awesome host! **Fran**, **Nishanth**, **Akesh** and **Mike**, for a good chat about various aspect of Ig genetics and analysis.

To **Larry Shapiro, Zizhang** and **Chaim** at Columbia University, for valuable discussion about the data analysis.

To the best corridor in MTC/Biomedicum:

Mikael Karlsson and **Lisa Westerberg**, for your B cell expertise and sharing the corridor/quarter with us. **Amanda**, for good occasional chats in the corridor and for being a great company at pubs, an additional last round of drinks and bottoms up! **Thomas**, for your cool attitude, innovative high-fives, giving me snus experience and for a great fun at the MTC pubs. **Silke**, for helping at the PhD registration seminars as a student representative. **Manasa**, for ‘how is it going?’ and taking care of my headphones at the parties. **Vanessa**, for being a good sport and good laughs. **Chenfei**, for friendliness. **Joanna and Milind**, for being awesome hosts and great parties. **Minghui**, for friendly smiles. **Mariana**, for fun at the course lab. **Marton**, for your cheerfulness. **Nikolai**, for the candies.

To the people at MTC:

Former head **Marie Arsenian Henriksson** and the present head **Pontus Aspenström** for a great working environment at the department. **Åsa** and **Gesan**, for always taking time to resolve queries and providing kind support and help related to doctoral education formalities at KI. **Kristina**, for your help with reimbursements. Former and present HR department: **Lina**, one of the best HR personnel I have met. **Helene** and **Martin**, for providing good administrative support. **Torbjörn, David, Christopher, Petra, and Susanne** in the Service group; for helping with the packages. **Per**, for helping with all IT related queries and the orders. **Magnus**, for providing technical support.

Kiran, thanks for your expert help and advices with sorting our precious NHP samples, good chats about life/science/career during long experiments, and for taking good care of the FACS Core Facility.

To the members of MTC student association and the PhD registration committee, thanks for all the meetings and discussion, it was a great learning experience

Friends and Family:

Monali, for your positivity and cheerfulness. Thank you for your friendship, all the help and fun moments over the years. Also, for being a worthy competitor in banter! **Gargi**, for making me feel calm, I feel connected to you. **Nilesh**, for all your help in critical times, for your kind nature and your passion for cricket. **Daya**, you are by far the most easy-going person I have met in Stockholm, I truly value your friendship, Thank you! **Sneha**, for your humble and caring nature. You are one of the most warm-hearted person among all my friends in Sweden. Your helpful nature is admirable. **Sachin T**, for your calmness and your love for yoga and photography. **Neha**, thank you for your selflessness and friendship, for fantastic dinners, movie nights, breakfasts and

offer/discount alerts. I feel at home with you all. **Daksh**, for reminding me how I used to be when I was kid. **Narayanan**, for being a caring person and for great help in crucial times – Thank you! Your energy is contagious, it is true pleasure to have you around, and five stars to Nara’s kitchen! **Ashish**, for nurturing our friendship, always being ready to help and for reminding me about life outside the lab, **Suhas**, thanks for breaking the ice (and I’m sorry for not doing so), for all the discussions on everything during lunch and coffee breaks, **Ashwini**, for your sincerity and smartness. Also, for surviving to my hilarity and for an awesome time on Norway trip. **Prajakta**, for occasional naivety and discussions about the work/life and for your amazing dance skills – like watching you dance! **Swapnali**, for making me comfortable to use (and not hesitating to learn) profane language during social bites. Your laughter confirms the success of my joke. **Sachin S**, for sharing the similar humour sense, your impeccable dance moves anytime on any track puts life in parties. **Ram**, for your time-saving python scripts, which has been very helpful during early days of data analysis – Thank you! **Ankur**, for your laugh and being a good sport to trolling, **Anuj**, for being that college ka कमीना दोस्त and for making me almost germfree human with your Manchurian gravy, (my gut flora is still recovering..). **Saranya**, for ‘The_She_Series’ and her perspective, for introducing the fun werewolf game to us. **Madhu**, for your hard work in the lab and in the basement (ping-pong) of MTC, **Ankit and Ishani**, for being a lovely couple and for complimenting my research work, **Shahul**, for your badminton and photography skills and tips, **Aarti**, for the character pillow and **Abhay**, for your wishes on festivals, **Suvarna and Deepak**, for being a good example of modern day parents, these skills have turned **Aarush** into a smart kid, **Gayatri and Vivek**, for being great fun at the parties, **Kunal, Abdul, Satendra**, for helping me to improve my badminton skills. **Chinmay**, for useful information about everything, **Amol**, for the Copenhagen trip, **Gaurav**, for your cleverness, bakar-gyan and for the pub plans, **Namita**, for being a great and kind host. **Sandip, Neha, Ranjana** for being nice hosts and the fun time in Lund. **Aditi**, for your help and for taking good care of me in Pune and **Savita**, for your greetings and compliments. All the best to both for your PhD defense. **Manisha, Shailesh, Gaurav B, Ketan, Sachin W, Vivek B, and Sonal N** - thanks very much for all your help and support during the admission and administrative processes before travelling to Sweden. **Girish and Praful**, for your help during initial days in Sweden. My best buddies **Vishal, Sneshank, and Ayaz**, for always being there for me. **Ajit and Nilesh K**, for all the encouragement during the early days of my career.

Family: **Dashrath** mama, **Suman** mavshi, **Vinod** jiju, and **Pandhari** dada, thank you for your constant help, support and encouragement. My cousins, **Sachin, Sushma, Rani, Pratik, and Priyanka** – you all are awesome! My nieces, **Jiya, Aarya, and Aaradhya** and nephew, **Devang** – I love you all so much! **Sonali**, for your trust, love, care and support! ☺

आई-बाबा - माझं प्रत्येक यश तुम्हाला समर्पित आहे. ☺

अंतिमतः मी हा शोधनिबंध माझ्या आयुष्यातील तीन महत्वाच्या व्यक्तींना समर्पित करू इच्छितो - माझी आई (मीरा), माझी शिक्षिका (गुणीला) आणि माझा मित्र (मनिषा ताई).

This thesis has been conducted within the Infection Biology program at the Department of Microbiology, Tumor and Cell Biology at Karolinska Institutet. The projects have been supported with funds from Karolinska Institutet (KID), The Swedish Research Council, International AIDS Vaccine Initiative (IAVI), and the National Institutes of Health (NIH).

9 REFERENCES

1. Fenner, F., *The Eradication of Smallpox*. Impact of Science on Society, 1988. **38**(2): p. 147-158.
2. L, P., *Méthode pour prévenir la rage après morsure*. Comptes Rendus des Séances de l'Académie des Sciences Paris, 1885(101): p. 765–772.
3. Plotkin, S., *History of vaccination*. Proc Natl Acad Sci U S A, 2014. **111**(34): p. 12283-7.
4. Plotkin, S.A. and S.L. Plotkin, *The development of vaccines: how the past led to the future*. Nat Rev Microbiol, 2011. **9**(12): p. 889-93.
5. Plotkin, S.A., *Vaccines: past, present and future*. Nat Med, 2005. **11**(4 Suppl): p. S5-11.
6. Koff, W.C., et al., *Accelerating next-generation vaccine development for global disease prevention*. Science, 2013. **340**(6136): p. 1232910.
7. Administration, U.F.a.D. *Vaccines, Blood & Biologics*. 2015; Available from: <http://www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/ucm093833.htm>.
8. Plotkin, S.A., *Correlates of protection induced by vaccination*. Clin Vaccine Immunol, 2010. **17**(7): p. 1055-65.
9. Pulendran, B. and R. Ahmed, *Immunological mechanisms of vaccination*. Nat Immunol, 2011. **12**(6): p. 509-17.
10. Amanna, I.J., N.E. Carlson, and M.K. Slifka, *Duration of humoral immunity to common viral and vaccine antigens*. N Engl J Med, 2007. **357**(19): p. 1903-15.
11. Bonsignori, M., et al., *HIV-1 envelope induces memory B cell responses that correlate with plasma antibody levels after envelope gp120 protein vaccination or HIV-1 infection*. J Immunol, 2009. **183**(4): p. 2708-17.
12. Xiang, Z., et al., *FcγRIIb controls bone marrow plasma cell persistence and apoptosis*. Nat Immunol, 2007. **8**(4): p. 419-29.
13. Radbruch, A., et al., *Competence and competition: the challenge of becoming a long-lived plasma cell*. Nat Rev Immunol, 2006. **6**(10): p. 741-50.
14. Odendahl, M., et al., *Generation of migratory antigen-specific plasma blasts and mobilization of resident plasma cells in a secondary immune response*. Blood, 2005. **105**(4): p. 1614-21.
15. Amanna, I.J. and M.K. Slifka, *Mechanisms that determine plasma cell lifespan and the duration of humoral immunity*. Immunol Rev, 2010. **236**: p. 125-38.
16. Slifka, M.K. and I. Amanna, *How advances in immunology provide insight into improving vaccine efficacy*. Vaccine, 2014. **32**(25): p. 2948-57.
17. Billips, L.G., et al., *Human B-cell development*. Ann N Y Acad Sci, 1995. **764**: p. 1-8.
18. Wardemann, H., et al., *Predominant autoantibody production by early human B cell precursors*. Science, 2003. **301**(5638): p. 1374-7.
19. Schroeder, H.W., Jr. and L. Cavacini, *Structure and function of immunoglobulins*. J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S41-52.
20. North, B., A. Lehmann, and R.L. Dunbrack, Jr., *A new clustering of antibody CDR loop conformations*. J Mol Biol, 2011. **406**(2): p. 228-56.
21. Xu, J.L. and M.M. Davis, *Diversity in the CDR3 region of V(H) is sufficient for most antibody specificities*. Immunity, 2000. **13**(1): p. 37-45.
22. Ippolito, G.C., et al., *Forced usage of positively charged amino acids in immunoglobulin CDR-H3 impairs B cell development and antibody production*. J Exp Med, 2006. **203**(6): p. 1567-78.

23. Matsuda, F., et al., *The complete nucleotide sequence of the human immunoglobulin heavy chain variable region locus*. J Exp Med, 1998. **188**(11): p. 2151-62.
24. Malcolm, S., et al., *Localization of human immunoglobulin kappa light chain variable region genes to the short arm of chromosome 2 by in situ hybridization*. Proc Natl Acad Sci U S A, 1982. **79**(16): p. 4957-61.
25. McBride, O.W., et al., *Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes*. J Exp Med, 1982. **155**(5): p. 1480-90.
26. Zachau, H.G., *The Human-Immunoglobulin Kappa Genes*. Progress in Immunology, Vol 7, 1989: p. 92-98.
27. Erikson, J., J. Martinis, and C.M. Croce, *Assignment of the genes for human lambda immunoglobulin chains to chromosome 22*. Nature, 1981. **294**(5837): p. 173-5.
28. Emanuel, B.S., et al., *Chromosomal orientation of the lambda light chain locus: V lambda is proximal to C lambda in 22q11*. Nucleic Acids Res, 1985. **13**(2): p. 381-7.
29. Kawasaki, K., et al., *One-megabase sequence analysis of the human immunoglobulin lambda gene locus*. Genome Research, 1997. **7**(3): p. 250-261.
30. Schroeder, H.W. and L. Cavacini, *Structure and function of immunoglobulins*. Journal of Allergy and Clinical Immunology, 2010. **125**(2): p. S41-S52.
31. Rhesus Macaque Genome, S., et al., *Evolutionary and biomedical insights from the rhesus macaque genome*. Science, 2007. **316**(5822): p. 222-34.
32. Sundling, C., et al., *High-resolution definition of vaccine-elicited B cell responses against the HIV primary receptor binding site*. Sci Transl Med, 2012. **4**(142): p. 142ra96.
33. Ramesh, A., et al., *Structure and Diversity of the Rhesus Macaque Immunoglobulin Loci through Multiple De Novo Genome Assemblies*. Front Immunol, 2017. **8**: p. 1407.
34. Watson, C.T. and F. Breden, *The immunoglobulin heavy chain locus: genetic variation, missing data, and implications for human disease*. Genes Immun, 2012. **13**(5): p. 363-73.
35. Watson, C.T., J. Glanville, and W.A. Marasco, *The Individual and Population Genetics of Antibody Immunity*. Trends Immunol, 2017. **38**(7): p. 459-470.
36. Fukui, K., et al., *Origin of adult T-cell leukemia virus. Implication for its zoonosis*. Mol Biol Med, 1983. **1**(4): p. 447-56.
37. Davies, J.M., T.A. Platts-Mills, and R.C. Aalberse, *The enigma of IgE+ B-cell memory in human subjects*. J Allergy Clin Immunol, 2013. **131**(4): p. 972-6.
38. Matthyssens, G. and T.H. Rabbitts, *Structure and multiplicity of genes for the human immunoglobulin heavy chain variable region*. Proc Natl Acad Sci U S A, 1980. **77**(11): p. 6561-5.
39. Walter, M.A., P.S. Linsley, and D.W. Cox, *Apa I polymorphism of a human immunoglobulin VH3 subclass locus*. Nucleic Acids Res, 1987. **15**(11): p. 4697.
40. Pincus, S.H., *Human immunoglobulin heavy-chain-variable (VH) region gene families defined by hybridization with cloned human and murine VH-genes*. Hum Immunol, 1988. **22**(3): p. 199-215.
41. Walter, M.A. and D.W. Cox, *Analysis of genetic variation reveals human immunoglobulin VH-region gene organization*. Am J Hum Genet, 1988. **42**(3): p. 446-51.
42. Sasso, E.H., K.W. Van Dijk, and E.C. Milner, *Prevalence and polymorphism of human VH3 genes*. J Immunol, 1990. **145**(8): p. 2751-7.

43. van Dijk, K.W., E.H. Sasso, and E.C. Milner, *Polymorphism of the human Ig VH4 gene family*. J Immunol, 1991. **146**(10): p. 3646-51.
44. Willems van Dijk, K., et al., *Heterogeneity in the human Ig VH locus*. J Immunol, 1989. **142**(7): p. 2547-54.
45. Adderson, E.E., et al., *The human VH3b gene subfamily is highly polymorphic*. J Immunol, 1993. **151**(2): p. 800-9.
46. Weng, N.P., et al., *Polymorphism of human immunoglobulin VH4 germ-line genes*. Eur J Immunol, 1992. **22**(4): p. 1075-82.
47. Scheepers, C., et al., *Ability to develop broadly neutralizing HIV-1 antibodies is not restricted by the germline Ig gene repertoire*. J Immunol, 2015. **194**(9): p. 4371-8.
48. Barbie, V. and M.P. Lefranc, *The human immunoglobulin kappa variable (IGKV) genes and joining (IGKJ) segments*. Exp Clin Immunogenet, 1998. **15**(3): p. 171-83.
49. Lefranc, M.P., *Nomenclature of the human immunoglobulin heavy (IGH) genes*. Exp Clin Immunogenet, 2001. **18**(2): p. 100-16.
50. Lefranc, M.P., et al., *IMGT(R), the international ImMunoGeneTics information system(R) 25 years on*. Nucleic Acids Res, 2015. **43**(Database issue): p. D413-22.
51. Pallares, N., et al., *The human immunoglobulin lambda variable (IGLV) genes and joining (IGLJ) segments*. Exp Clin Immunogenet, 1998. **15**(1): p. 8-18.
52. Pallares, N., et al., *The human immunoglobulin heavy variable genes*. Exp Clin Immunogenet, 1999. **16**(1): p. 36-60.
53. Wang, Y., et al., *Many human immunoglobulin heavy-chain IGHV gene polymorphisms have been reported in error*. Immunol Cell Biol, 2008. **86**(2): p. 111-5.
54. Karlsson Hedestam, G.B., et al., *Evolution of B cell analysis and Env trimer redesign*. Immunol Rev, 2017. **275**(1): p. 183-202.
55. Watson, C.T., et al., *Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation*. Am J Hum Genet, 2013. **92**(4): p. 530-46.
56. Gadala-Maria, D., et al., *Automated analysis of high-throughput B-cell sequencing data reveals a high frequency of novel immunoglobulin V gene segment alleles*. Proc Natl Acad Sci U S A, 2015. **112**(8): p. E862-70.
57. Collins, A.M., et al., *The mouse antibody heavy chain repertoire is germline-focused and highly variable between inbred strains*. Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1676).
58. Kirik, U., et al., *Parallel antibody germline gene and haplotype analyses support the validity of immunoglobulin germline gene inference and discovery*. Mol Immunol, 2017. **87**: p. 12-22.
59. Corcoran, M.M., et al., *Production of individualized V gene databases reveals high levels of immunoglobulin genetic diversity*. Nat Commun, 2016. **7**: p. 13642.
60. Janeway, C.A., et al., *Immunobiology*. Fifth edition ed. 2001: Curent Biology Publications.
61. Lucas, J.S., et al., *Chapter 1 - The Structure and Regulation of the Immunoglobulin Loci*, in *Molecular Biology of B Cells (Second Edition)*, F.W.A.H.R. Reth, Editor. 2015, Academic Press: London. p. 1-11.
62. Little, A.J., et al., *Chapter 2 - The Mechanism of V(D)J Recombination*, in *Molecular Biology of B Cells (Second Edition)*, F.W.A.H.R. Reth, Editor. 2015, Academic Press: London. p. 13-34.

63. Di Niro, R., et al., *Salmonella Infection Drives Promiscuous B Cell Activation Followed by Extrafollicular Affinity Maturation*. Immunity, 2015. **43**(1): p. 120-131.
64. Chang, B. and P. Casali, *The CDR1 sequences of a major proportion of human germline Ig VH genes are inherently susceptible to amino acid replacement*. Immunol Today, 1994. **15**(8): p. 367-73.
65. Milstein, C., M.S. Neuberger, and R. Staden, *Both DNA strands of antibody genes are hypermutation targets*. Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8791-4.
66. Wu, X., et al., *Immunoglobulin somatic hypermutation: double-strand DNA breaks, AID and error-prone DNA repair*. J Clin Immunol, 2003. **23**(4): p. 235-46.
67. Li, Z., et al., *The generation of antibody diversity through somatic hypermutation and class switch recombination*. Genes Dev, 2004. **18**(1): p. 1-11.
68. Xu, Z., et al., *DNA lesions and repair in immunoglobulin class switch recombination and somatic hypermutation*. Ann N Y Acad Sci, 2005. **1050**: p. 146-62.
69. Longerich, S., et al., *The very 5' end and the constant region of Ig genes are spared from somatic mutation because AID does not access these regions*. J Exp Med, 2005. **202**(10): p. 1443-54.
70. Rada, C. and C. Milstein, *The intrinsic hypermutability of antibody heavy and light chain genes decays exponentially*. EMBO J, 2001. **20**(16): p. 4570-6.
71. Pape, K.A., et al., *Visualization of the genesis and fate of isotype-switched B cells during a primary immune response*. J Exp Med, 2003. **197**(12): p. 1677-87.
72. Bergqvist, P., et al., *Gut IgA class switch recombination in the absence of CD40 does not occur in the lamina propria and is independent of germinal centers*. J Immunol, 2006. **177**(11): p. 7772-83.
73. Kunkel, E.J. and E.C. Butcher, *Plasma-cell homing*. Nat Rev Immunol, 2003. **3**(10): p. 822-9.
74. Ellyard, J.I., et al., *Contribution of stromal cells to the migration, function and retention of plasma cells in human spleen: potential roles of CXCL12, IL-6 and CD54*. Eur J Immunol, 2005. **35**(3): p. 699-708.
75. Ellyard, J.I., et al., *Antigen-selected, immunoglobulin-secreting cells persist in human spleen and bone marrow*. Blood, 2004. **103**(10): p. 3805-12.
76. Winter, O., et al., *Megakaryocytes constitute a functional component of a plasma cell niche in the bone marrow*. Blood, 2010. **116**(11): p. 1867-75.
77. Victora, G.D. and M.C. Nussenzweig, *Germinal centers*. Annu Rev Immunol, 2012. **30**: p. 429-57.
78. Shaffer, A.L., et al., *BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control*. Immunity, 2000. **13**(2): p. 199-212.
79. McHeyzer-Williams, M., et al., *Molecular programming of B cell memory*. Nature Reviews Immunology, 2012. **12**(1): p. 24-34.
80. Shih, T.A.Y., et al., *Role of BCR affinity in T cell-dependent antibody responses in vivo*. Nature Immunology, 2002. **3**(6): p. 570-575.
81. Dal Porto, J.M., et al., *Very low affinity B cells form germinal centers, become memory B cells, and participate in secondary immune responses when higher affinity competition is reduced*. Journal of Experimental Medicine, 2002. **195**(9): p. 1215-1221.
82. Barnett, B., et al., *Asymmetric division of germinal center B cells*. Journal of Immunology, 2012. **188**.

83. Barnett, B.E., et al., *Asymmetric B Cell Division in the Germinal Center Reaction*. Science, 2012. **335**(6066): p. 342-344.
84. Liu, Y.J., S. Oldfield, and I.C.M. Maclellan, *Memory B-Cells in T-Cell-Dependent Antibody-Responses Colonize the Splenic Marginal Zones*. European Journal of Immunology, 1988. **18**(3): p. 355-362.
85. Dunnwalters, D.K., P.G. Isaacson, and J. Spencer, *Analysis of Mutations in Immunoglobulin Heavy-Chain Variable Region Genes of Microdissected Marginal Zone (Mgz) B-Cells Suggests That the Mgz of Human Spleen Is a Reservoir of Memory B-Cells*. Journal of Experimental Medicine, 1995. **182**(2): p. 559-566.
86. Ehrhardt, G.R.A., et al., *Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells*. Journal of Experimental Medicine, 2005. **202**(6): p. 783-791.
87. Aiba, Y., et al., *Preferential localization of IgG memory B cells adjacent to contracted germinal centers*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(27): p. 12192-12197.
88. Crotty, S., et al., *Cutting edge: Long-term B cell memory in humans after smallpox vaccination*. Journal of Immunology, 2003. **171**(10): p. 4969-4973.
89. Lee, F.E.H., et al., *Circulating Human Antibody-Secreting Cells during Vaccinations and Respiratory Viral Infections Are Characterized by High Specificity and Lack of Bystander Effect*. Journal of Immunology, 2011. **186**(9): p. 5514-5521.
90. Benson, M.J., et al., *Distinction of the memory B cell response to cognate antigen versus bystander inflammatory signals*. J Exp Med, 2009. **206**(9): p. 2013-25.
91. Ahmed, R. and D. Gray, *Immunological memory and protective immunity: understanding their relation*. Science, 1996. **272**(5258): p. 54-60.
92. Rajewsky, K., *Clonal selection and learning in the antibody system*. Nature, 1996. **381**(6585): p. 751-8.
93. Dogan, I., et al., *Multiple layers of B cell memory with different effector functions*. Nat Immunol, 2009. **10**(12): p. 1292-9.
94. Batista, F.D. and N.E. Harwood, *The who, how and where of antigen presentation to B cells*. Nat Rev Immunol, 2009. **9**(1): p. 15-27.
95. Dorner, T. and A. Radbruch, *Selecting B cells and plasma cells to memory*. J Exp Med, 2005. **201**(4): p. 497-9.
96. Tokoyoda, K., et al., *Cellular niches controlling B lymphocyte behavior within bone marrow during development*. Immunity, 2004. **20**(6): p. 707-18.
97. Nakayama, T., et al., *Cutting edge: profile of chemokine receptor expression on human plasma cells accounts for their efficient recruitment to target tissues*. J Immunol, 2003. **170**(3): p. 1136-40.
98. Peperzak, V., et al., *Mcl-1 is essential for the survival of plasma cells*. Nat Immunol, 2013. **14**(3): p. 290-7.
99. Ahuja, A., et al., *Maintenance of the plasma cell pool is independent of memory B cells*. Proc Natl Acad Sci U S A, 2008. **105**(12): p. 4802-7.
100. Slifka, M.K. and R. Ahmed, *Long-lived plasma cells: a mechanism for maintaining persistent antibody production*. Curr Opin Immunol, 1998. **10**(3): p. 252-8.
101. Manz, R.A., et al., *Survival of long-lived plasma cells is independent of antigen*. Int Immunol, 1998. **10**(11): p. 1703-11.
102. Hammarlund, E., et al., *Plasma cell survival in the absence of B cell memory*. Nat Commun, 2017. **8**(1): p. 1781.

103. Halliley, J.L., et al., *Long-Lived Plasma Cells Are Contained within the CD19(-)CD38(hi)CD138(+) Subset in Human Bone Marrow*. *Immunity*, 2015. **43**(1): p. 132-45.
104. McCoy, L.E. and D.R. Burton, *Identification and specificity of broadly neutralizing antibodies against HIV*. *Immunol Rev*, 2017. **275**(1): p. 11-20.
105. Soldemo, M. and G.B. Karlsson Hedestam, *Env-Specific Antibodies in Chronic Infection versus in Vaccination*. *Front Immunol*, 2017. **8**: p. 1057.
106. Vollmers, C., et al., *Genetic measurement of memory B-cell recall using antibody repertoire sequencing*. *Proc Natl Acad Sci U S A*, 2013. **110**(33): p. 13463-8.
107. Sakakibara, S., et al., *Clonal evolution and antigen recognition of anti-nuclear antibodies in acute systemic lupus erythematosus*. *Sci Rep*, 2017. **7**(1): p. 16428.
108. Amara, K., et al., *Monoclonal IgG antibodies generated from joint-derived B cells of RA patients have a strong bias toward citrullinated autoantigen recognition*. *J Exp Med*, 2013. **210**(3): p. 445-55.
109. Meffre, E. and H. Wardemann, *B-cell tolerance checkpoints in health and autoimmunity*. *Curr Opin Immunol*, 2008. **20**(6): p. 632-8.
110. Scheid, J.F., et al., *Differential regulation of self-reactivity discriminates between IgG⁺ human circulating memory B cells and bone marrow plasma cells*. *Proc Natl Acad Sci U S A*, 2011. **108**(44): p. 18044-8.
111. Wrammert, J., et al., *Rapid cloning of high-affinity human monoclonal antibodies against influenza virus*. *Nature*, 2008. **453**(7195): p. 667-71.
112. Corti, D. and A. Lanzavecchia, *Broadly neutralizing antiviral antibodies*. *Annu Rev Immunol*, 2013. **31**: p. 705-42.
113. Fu, Y., et al., *A broadly neutralizing anti-influenza antibody reveals ongoing capacity of haemagglutinin-specific memory B cells to evolve*. *Nat Commun*, 2016. **7**: p. 12780.
114. Tan, J., et al., *A LAIR1 insertion generates broadly reactive antibodies against malaria variant antigens*. *Nature*, 2016. **529**(7584): p. 105-109.
115. Dejnirattisai, W., et al., *A new class of highly potent, broadly neutralizing antibodies isolated from viremic patients infected with dengue virus*. *Nat Immunol*, 2015. **16**(2): p. 170-177.
116. Robbiani, D.F., et al., *Recurrent Potent Human Neutralizing Antibodies to Zika Virus in Brazil and Mexico*. *Cell*, 2017. **169**(4): p. 597-609 e11.
117. Wu, X., et al., *Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing*. *Science*, 2011. **333**(6049): p. 1593-602.
118. Doria-Rose, N.A., et al., *Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies*. *Nature*, 2014. **509**(7498): p. 55-62.
119. Sanders, R.W. and J.P. Moore, *Native-like Env trimers as a platform for HIV-1 vaccine design*. *Immunol Rev*, 2017. **275**(1): p. 161-182.
120. Sanders, R.W., et al., *HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers*. *Science*, 2015. **349**(6244): p. aac4223.
121. Martinez-Murillo, P., et al., *Particulate Array of Well-Ordered HIV Clade C Env Trimers Elicits Neutralizing Antibodies that Display a Unique V2 Cap Approach*. *Immunity*, 2017. **46**(5): p. 804-817 e7.
122. Phad, G.E., et al., *Diverse antibody genetic and recognition properties revealed following HIV-1 envelope glycoprotein immunization*. *J Immunol*, 2015. **194**(12): p. 5903-14.
123. Sundling, C., et al., *Single-cell and deep sequencing of IgG-switched macaque B cells reveal a diverse Ig repertoire following immunization*. *J Immunol*, 2014. **192**(8): p. 3637-44.

124. Tran, K., et al., *Vaccine-elicited primate antibodies use a distinct approach to the HIV-1 primary receptor binding site informing vaccine redesign*. Proc Natl Acad Sci U S A, 2014. **111**(7): p. E738-47.
125. Wang, Y., et al., *High-Resolution Longitudinal Study of HIV-1 Env Vaccine-Elicited B Cell Responses to the Virus Primary Receptor Binding Site Reveals Affinity Maturation and Clonal Persistence*. J Immunol, 2016. **196**(9): p. 3729-43.
126. Reuter, J.A., D.V. Spacek, and M.P. Snyder, *High-throughput sequencing technologies*. Mol Cell, 2015. **58**(4): p. 586-97.
127. Boyd, S.D. and S.A. Joshi, *High-Throughput DNA Sequencing Analysis of Antibody Repertoires*. Microbiol Spectr, 2014. **2**(5).
128. Kovaltsuk, A., et al., *How B-Cell Receptor Repertoire Sequencing Can Be Enriched with Structural Antibody Data*. Front Immunol, 2017. **8**: p. 1753.
129. Hou, D., et al., *High-Throughput Sequencing-Based Immune Repertoire Study during Infectious Disease*. Front Immunol, 2016. **7**: p. 336.
130. Larsen, P.A. and T.P. Smith, *Application of circular consensus sequencing and network analysis to characterize the bovine IgG repertoire*. BMC Immunol, 2012. **13**: p. 52.
131. He, L., et al., *Toward a more accurate view of human B-cell repertoire by next-generation sequencing, unbiased repertoire capture and single-molecule barcoding*. Sci Rep, 2014. **4**: p. 6778.
132. Quail, M.A., et al., *A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers*. BMC Genomics, 2012. **13**: p. 341.
133. Apostolaei, A.J.T., J.R. Review, *Synthesis, and Application of Information on the Human Lymphatic System to Radiation Dosimetry for Chronic Lymphocytic Leukemia*. SENES Oak Ridge, Inc., Oak Ridge, TN, 2010.
134. Collins, A.M. and K.J.L. Jackson, *On being the right size: antibody repertoire formation in the mouse and human*. Immunogenetics, 2018. **70**(3): p. 143-158.
135. Steiniger, B.S., *Human spleen microanatomy: why mice do not suffice*. Immunology, 2015. **145**(3): p. 334-346.
136. Benitez, A., et al., *Differences in Mouse and Human Nonmemory B Cell Pools*. Journal of Immunology, 2014. **192**(10): p. 4610-4619.
137. Georgiou, G., et al., *The promise and challenge of high-throughput sequencing of the antibody repertoire*. Nature Biotechnology, 2014. **32**(2): p. 158-168.
138. Wardemann, H. and C.E. Busse, *Novel Approaches to Analyze Immunoglobulin Repertoires*. Trends in Immunology, 2017. **38**(7): p. 471-482.
139. Best, K., et al., *Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding*. Sci Rep, 2015. **5**: p. 14629.
140. Boyd, S.D. and J.E. Crowe, *Deep sequencing and human antibody repertoire analysis*. Current Opinion in Immunology, 2016. **40**: p. 103-109.
141. Warren, R.L., et al., *Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes*. Genome Research, 2011. **21**(5): p. 790-797.
142. Ferrarini, M. and M. Chiorazzi, *Recent advances in the molecular biology and immunobiology of chronic lymphocytic leukemia*. Seminars in Hematology, 2004. **41**(3): p. 207-223.
143. Varadarajan, N., et al., *A high-throughput single-cell analysis of human CD8(+) T cell functions reveals discordance for cytokine secretion and cytotoxicity*. Journal of Clinical Investigation, 2011. **121**(11): p. 4322-4331.

144. Greiff, V., et al., *Bioinformatic and Statistical Analysis of Adaptive Immune Repertoires*. Trends in Immunology, 2015. **36**(11): p. 738-749.
145. Breden, F., et al., *Reproducibility and Reuse of Adaptive Immune Receptor Repertoire Data*. Front Immunol, 2017. **8**: p. 1418.
146. Rubelt, F., et al., *Adaptive Immune Receptor Repertoire Community recommendations for sharing immune-repertoire sequencing data*. Nat Immunol, 2017. **18**(12): p. 1274-1278.
147. Miho, E., et al., *Computational Strategies for Dissecting the High-Dimensional Complexity of Adaptive Immune Repertoires*. Front Immunol, 2018. **9**: p. 224.
148. Menzel, U., et al., *Comprehensive evaluation and optimization of amplicon library preparation methods for high-throughput antibody sequencing*. PLoS One, 2014. **9**(5): p. e96727.
149. Ye, J., et al., *IgBLAST: an immunoglobulin variable domain sequence analysis tool*. Nucleic Acids Res, 2013. **41**(Web Server issue): p. W34-40.
150. Gaeta, B.A., et al., *iHMMune-align: hidden Markov model-based alignment and identification of germline genes in rearranged immunoglobulin gene sequences*. Bioinformatics, 2007. **23**(13): p. 1580-7.
151. Bolotin, D.A., et al., *MiXCR: software for comprehensive adaptive immunity profiling*. Nat Methods, 2015. **12**(5): p. 380-1.
152. Yaari, G. and S.H. Kleinstein, *Practical guidelines for B-cell receptor repertoire sequencing analysis*. Genome Med, 2015. **7**: p. 121.
153. Hershberg, U. and E.T. Luning Prak, *The analysis of clonal expansions in normal and autoimmune B cell repertoires*. Philos Trans R Soc Lond B Biol Sci, 2015. **370**(1676).
154. Tipton, C.M., et al., *Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus*. Nat Immunol, 2015. **16**(7): p. 755-65.
155. Madi, A., et al., *T-cell receptor repertoires share a restricted set of public and abundant CDR3 sequences that are associated with self-related immunity*. Genome Res, 2014. **24**(10): p. 1603-12.
156. Greiff, V., et al., *Quantitative assessment of the robustness of next-generation sequencing of antibody variable gene repertoires from immunized mice*. BMC Immunol, 2014. **15**: p. 40.
157. Wang, C., et al., *B-cell repertoire responses to varicella-zoster vaccination in human identical twins*. Proc Natl Acad Sci U S A, 2015. **112**(2): p. 500-5.
158. UNAIDS. 2014; Available from: <http://www.unaids.org/en/resources/campaigns/HowAIDSchangedeverything/factsheet>.
159. Frankel, A.D. and J.A. Young, *HIV-1: fifteen proteins and an RNA*. Annu Rev Biochem, 1998. **67**: p. 1-25.
160. Vaishnav, Y.N. and F. Wong-Staal, *The biochemistry of AIDS*. Annu Rev Biochem, 1991. **60**: p. 577-630.
161. Wyatt, R. and J. Sodroski, *The HIV-1 envelope glycoproteins: fusogens, antigens, and immunogens*. Science, 1998. **280**(5371): p. 1884-8.
162. Stein, B.S. and E.G. Engleman, *Intracellular processing of the gp160 HIV-1 envelope precursor. Endoproteolytic cleavage occurs in a cis or medial compartment of the Golgi complex*. J Biol Chem, 1990. **265**(5): p. 2640-9.
163. Earl, P.L., B. Moss, and R.W. Doms, *Folding, interaction with GRP78-BiP, assembly, and transport of the human immunodeficiency virus type 1 envelope protein*. J Virol, 1991. **65**(4): p. 2047-55.

164. Earl, P.L., R.W. Doms, and B. Moss, *Oligomeric structure of the human immunodeficiency virus type 1 envelope glycoprotein*. Proc Natl Acad Sci U S A, 1990. **87**(2): p. 648-52.
165. Leonard, C.K., et al., *Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells*. J Biol Chem, 1990. **265**(18): p. 10373-82.
166. Chan, D.C., et al., *Core structure of gp41 from the HIV envelope glycoprotein*. Cell, 1997. **89**(2): p. 263-73.
167. Weissenhorn, W., et al., *Atomic structure of the ectodomain from HIV-1 gp41*. Nature, 1997. **387**(6631): p. 426-30.
168. Mizuochi, T., et al., *Diversity of oligosaccharide structures on the envelope glycoprotein gp 120 of human immunodeficiency virus 1 from the lymphoblastoid cell line H9. Presence of complex-type oligosaccharides with bisecting N-acetylglucosamine residues*. J Biol Chem, 1990. **265**(15): p. 8519-24.
169. Center, R.J., et al., *Oligomeric structure of the human immunodeficiency virus type 1 envelope protein on the virion surface*. J Virol, 2002. **76**(15): p. 7863-7.
170. Lu, M., S.C. Blacklow, and P.S. Kim, *A trimeric structural domain of the HIV-1 transmembrane glycoprotein*. Nat Struct Biol, 1995. **2**(12): p. 1075-82.
171. Horiya, S., I.S. MacPherson, and I.J. Krauss, *Recent strategies targeting HIV glycans in vaccine design*. Nat Chem Biol, 2014. **10**(12): p. 990-9.
172. Parren, P.W., D.R. Burton, and Q.J. Sattentau, *HIV-1 antibody--debris or virion?* Nat Med, 1997. **3**(4): p. 366-7.
173. Wyatt, R., et al., *Analysis of the interaction of the human immunodeficiency virus type 1 gp120 envelope glycoprotein with the gp41 transmembrane glycoprotein*. J Virol, 1997. **71**(12): p. 9722-31.
174. Moore, J.P. and J. Sodroski, *Antibody cross-competition analysis of the human immunodeficiency virus type 1 gp120 exterior envelope glycoprotein*. J Virol, 1996. **70**(3): p. 1863-72.
175. Zhu, P., et al., *Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions*. Proc Natl Acad Sci U S A, 2003. **100**(26): p. 15812-7.
176. Zhu, P., et al., *Distribution and three-dimensional structure of AIDS virus envelope spikes*. Nature, 2006. **441**(7095): p. 847-52.
177. Wyatt, R., et al., *The antigenic structure of the HIV gp120 envelope glycoprotein*. Nature, 1998. **393**(6686): p. 705-11.
178. Kwong, P.D., et al., *Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody*. Nature, 1998. **393**(6686): p. 648-59.
179. LaLonde, J.M., et al., *Structure-based design, synthesis, and characterization of dual hotspot small-molecule HIV-1 entry inhibitors*. J Med Chem, 2012. **55**(9): p. 4382-96.
180. Kwong, P.D., et al., *Structures of HIV-1 gp120 envelope glycoproteins from laboratory-adapted and primary isolates*. Structure, 2000. **8**(12): p. 1329-39.
181. Huang, C.C., et al., *Structure of a V3-containing HIV-1 gp120 core*. Science, 2005. **310**(5750): p. 1025-8.
182. Zhou, T., et al., *Structural definition of a conserved neutralization epitope on HIV-1 gp120*. Nature, 2007. **445**(7129): p. 732-7.
183. Pancera, M., et al., *Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility*. Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1166-71.

184. Chen, L., et al., *Structural basis of immune evasion at the site of CD4 attachment on HIV-1 gp120*. Science, 2009. **326**(5956): p. 1123-7.
185. Diskin, R., P.M. Marcovecchio, and P.J. Bjorkman, *Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity*. Nat Struct Mol Biol, 2010. **17**(5): p. 608-13.
186. Zhou, T., et al., *Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01*. Science, 2010. **329**(5993): p. 811-7.
187. Wu, X., et al., *Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing*. Science, 2011. **333**(6049): p. 1593-602.
188. Pejchal, R., et al., *A potent and broad neutralizing antibody recognizes and penetrates the HIV glycan shield*. Science, 2011. **334**(6059): p. 1097-103.
189. Diskin, R., et al., *Increasing the potency and breadth of an HIV antibody by using structure-based rational design*. Science, 2011. **334**(6060): p. 1289-93.
190. Kwon, Y.D., et al., *Unliganded HIV-1 gp120 core structures assume the CD4-bound conformation with regulation by quaternary interactions and variable loops*. Proc Natl Acad Sci U S A, 2012. **109**(15): p. 5663-8.
191. Wu, X., et al., *Neutralization escape variants of human immunodeficiency virus type 1 are transmitted from mother to infant*. J Virol, 2006. **80**(2): p. 835-44.
192. Sanders, R.W., et al., *A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies*. PLoS Pathog, 2013. **9**(9): p. e1003618.
193. Julien, J.P., et al., *Crystal structure of a soluble cleaved HIV-1 envelope trimer*. Science, 2013. **342**(6165): p. 1477-83.
194. Lyumkis, D., et al., *Cryo-EM structure of a fully glycosylated soluble cleaved HIV-1 envelope trimer*. Science, 2013. **342**(6165): p. 1484-90.
195. Bartesaghi, A., et al., *Prefusion structure of trimeric HIV-1 envelope glycoprotein determined by cryo-electron microscopy*. Nat Struct Mol Biol, 2013. **20**(12): p. 1352-7.
196. Guenaga, J., et al., *Glycine Substitution at Helix-to-Coil Transitions Facilitates the Structural Determination of a Stabilized Subtype C HIV Envelope Glycoprotein*. Immunity, 2017. **46**(5): p. 792-803 e3.
197. Poignard, P., et al., *Heterogeneity of envelope molecules expressed on primary human immunodeficiency virus type 1 particles as probed by the binding of neutralizing and nonneutralizing antibodies*. J Virol, 2003. **77**(1): p. 353-65.
198. Moore, P.L., et al., *Nature of nonfunctional envelope proteins on the surface of human immunodeficiency virus type 1*. J Virol, 2006. **80**(5): p. 2515-28.
199. Tomaras, G.D., et al., *Initial B Cell Responses to Transmitted HIV-1: Virion-Binding IgM and IgG Antibodies Followed by Plasma Anti-gp41 Antibodies with Ineffective Control of Initial Viremia*. J Virol, 2008.
200. Keele, B.F., et al., *Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection*. Proc Natl Acad Sci U S A, 2008. **105**(21): p. 7552-7.
201. Wei, X., et al., *Antibody neutralization and escape by HIV-1*. Nature, 2003. **422**(6929): p. 307-12.
202. Richman, D.D., et al., *Rapid evolution of the neutralizing antibody response to HIV type 1 infection*. Proc Natl Acad Sci U S A, 2003. **100**(7): p. 4144-9.
203. Aasa-Chapman, M.M., et al., *Development of the antibody response in acute HIV-1 infection*. AIDS, 2004. **18**(3): p. 371-81.
204. Gray, E.S., et al., *Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection*. J Virol, 2007. **81**(12): p. 6187-96.

205. Frost, S.D., et al., *Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection*. Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18514-9.
206. Overbaugh, J. and L.M. Rudensey, *Alterations in potential sites for glycosylation predominate during evolution of the simian immunodeficiency virus envelope gene in macaques*. J Virol, 1992. **66**(10): p. 5937-48.
207. Chackerian, B., L.M. Rudensey, and J. Overbaugh, *Specific N-linked and O-linked glycosylation modifications in the envelope V1 domain of simian immunodeficiency virus variants that evolve in the host alter recognition by neutralizing antibodies*. J Virol, 1997. **71**(10): p. 7719-27.
208. Bunnik, E.M., et al., *Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection*. J Virol, 2008. **82**(16): p. 7932-41.
209. Sather, D.N., et al., *Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection*. J Virol, 2009. **83**(2): p. 757-69.
210. Li, Y., et al., *Analysis of neutralization specificities in polyclonal sera derived from human immunodeficiency virus type 1-infected individuals*. J Virol, 2009. **83**(2): p. 1045-59.
211. Doria-Rose, N.A., et al., *Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies*. J Virol, 2009. **83**(1): p. 188-99.
212. Binley, J.M., et al., *Profiling the specificity of neutralizing antibodies in a large panel of plasmas from patients chronically infected with human immunodeficiency virus type 1 subtypes B and C*. J Virol, 2008. **82**(23): p. 11651-68.
213. Dhillon, A.K., et al., *Dissecting the neutralizing antibody specificities of broadly neutralizing sera from human immunodeficiency virus type 1-infected donors*. J Virol, 2007. **81**(12): p. 6548-62.
214. Doria-Rose, N.A., et al., *Breadth of human immunodeficiency virus-specific neutralizing activity in sera: clustering analysis and association with clinical variables*. J Virol, 2010. **84**(3): p. 1631-6.
215. Hraber, P., et al., *Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection*. AIDS, 2014. **28**(2): p. 163-9.
216. Moore, P.L., et al., *Potent and broad neutralization of HIV-1 subtype C by plasma antibodies targeting a quaternary epitope including residues in the V2 loop*. J Virol, 2011. **85**(7): p. 3128-41.
217. Simek, M.D., et al., *Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm*. J Virol, 2009. **83**(14): p. 7337-48.
218. Gray, E.S., et al., *Antibody specificities associated with neutralization breadth in plasma from human immunodeficiency virus type 1 subtype C-infected blood donors*. J Virol, 2009. **83**(17): p. 8925-37.
219. Goo, L., et al., *Early development of broadly neutralizing antibodies in HIV-1-infected infants*. Nat Med, 2014. **20**(6): p. 655-8.
220. Georgiev, I.S., et al., *Delineating antibody recognition in polyclonal sera from patterns of HIV-1 isolate neutralization*. Science, 2013. **340**(6133): p. 751-6.
221. Tomaras, G.D., et al., *Polyclonal B cell responses to conserved neutralization epitopes in a subset of HIV-1-infected individuals*. J Virol, 2011. **85**(21): p. 11502-19.

222. Walker, L.M., et al., *A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals*. PLoS Pathog, 2010. **6**(8): p. e1001028.
223. Klein, F., et al., *Antibodies in HIV-1 vaccine development and therapy*. Science, 2013. **341**(6151): p. 1199-204.
224. West, A.P., Jr., et al., *Structural insights on the role of antibodies in HIV-1 vaccine and therapy*. Cell, 2014. **156**(4): p. 633-48.
225. Klein, F., et al., *HIV therapy by a combination of broadly neutralizing antibodies in humanized mice*. Nature, 2012. **492**(7427): p. 118-22.
226. Barouch, D.H., et al., *Therapeutic efficacy of potent neutralizing HIV-1-specific monoclonal antibodies in SHIV-infected rhesus monkeys*. Nature, 2013. **503**(7475): p. 224-8.
227. Shingai, M., et al., *Antibody-mediated immunotherapy of macaques chronically infected with SHIV suppresses viraemia*. Nature, 2013. **503**(7475): p. 277-80.
228. Horwitz, J.A., et al., *HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16538-43.
229. Moldt, B., et al., *Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo*. Proc Natl Acad Sci U S A, 2012. **109**(46): p. 18921-5.
230. Kwong, P.D., et al., *Antibodyomics: bioinformatics technologies for understanding B-cell immunity to HIV-1*. Immunol Rev, 2017. **275**(1): p. 108-128.
231. Mascola, J.R. and B.F. Haynes, *HIV-1 neutralizing antibodies: understanding nature's pathways*. Immunol Rev, 2013. **254**(1): p. 225-44.
232. Kwong, P.D., J.R. Mascola, and G.J. Nabel, *Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning*. Nat Rev Immunol, 2013. **13**(9): p. 693-701.
233. Kwong, P.D. and J.R. Mascola, *Human antibodies that neutralize HIV-1: identification, structures, and B cell ontogenies*. Immunity, 2012. **37**(3): p. 412-25.
234. Burton, D.R. and L. Hangartner, *Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design*. Annu Rev Immunol, 2016. **34**: p. 635-59.
235. Esparza, J., *A brief history of the global effort to develop a preventive HIV vaccine*. Vaccine, 2013. **31**(35): p. 3502-18.
236. Gilbert, P.B., et al., *Correlation between immunologic responses to a recombinant glycoprotein 120 vaccine and incidence of HIV-1 infection in a phase 3 HIV-1 preventive vaccine trial*. J Infect Dis, 2005. **191**(5): p. 666-77.
237. Rerks-Ngarm, S., et al., *Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand*. N Engl J Med, 2009. **361**(23): p. 2209-20.
238. Forsell, M.N., W.R. Schief, and R.T. Wyatt, *Immunogenicity of HIV-1 envelope glycoprotein oligomers*. Curr Opin HIV AIDS, 2009. **4**(5): p. 380-7.
239. Yang, X., et al., *Characterization of stable, soluble trimers containing complete ectodomains of human immunodeficiency virus type 1 envelope glycoproteins*. J Virol, 2000. **74**(12): p. 5716-25.
240. Yang, X., et al., *Modifications that stabilize human immunodeficiency virus envelope glycoprotein trimers in solution*. J Virol, 2000. **74**(10): p. 4746-54.
241. Forsell, M.N., et al., *Biochemical and immunogenic characterization of soluble human immunodeficiency virus type 1 envelope glycoprotein trimers expressed by semliki forest virus*. J Virol, 2005. **79**(17): p. 10902-14.

242. Sundback, M., et al., *Efficient expansion of HIV-1-specific T cell responses by homologous immunization with recombinant Semliki Forest virus particles*. Virology, 2005. **341**(2): p. 190-202.
243. Li, Y., et al., *Characterization of antibody responses elicited by human immunodeficiency virus type 1 primary isolate trimeric and monomeric envelope glycoproteins in selected adjuvants*. J Virol, 2006. **80**(3): p. 1414-26.
244. Yang, X., R. Wyatt, and J. Sodroski, *Improved elicitation of neutralizing antibodies against primary human immunodeficiency viruses by soluble stabilized envelope glycoprotein trimers*. J Virol, 2001. **75**(3): p. 1165-71.
245. Sundling, C., et al., *Immunization of macaques with soluble HIV type 1 and influenza virus envelope glycoproteins results in a similarly rapid contraction of peripheral B-cell responses after boosting*. J Infect Dis, 2013. **207**(3): p. 426-31.
246. Douagi, I., et al., *Influence of novel CD4 binding-defective HIV-1 envelope glycoprotein immunogens on neutralizing antibody and T-cell responses in nonhuman primates*. J Virol, 2010. **84**(4): p. 1683-95.
247. Sundling, C., et al., *Immunization with wild-type or CD4-binding-defective HIV-1 Env trimers reduces viremia equivalently following heterologous challenge with simian-human immunodeficiency virus*. J Virol, 2010. **84**(18): p. 9086-95.
248. Sundling, C., et al., *Soluble HIV-1 Env trimers in adjuvant elicit potent and diverse functional B cell responses in primates*. J Exp Med, 2010. **207**(9): p. 2003-17.
249. Sharma, S.K., et al., *Cleavage-independent HIV-1 Env trimers engineered as soluble native spike mimetics for vaccine design*. Cell Rep, 2015. **11**(4): p. 539-50.
250. Guenaga, J., et al., *Well-ordered trimeric HIV-1 subtype B and C soluble spike mimetics generated by negative selection display native-like properties*. PLoS Pathog, 2015. **11**(1): p. e1004570.
251. Go, E.P., et al., *Native Conformation and Canonical Disulfide Bond Formation Are Interlinked Properties of HIV-1 Env Glycoproteins*. J Virol, 2015. **90**(6): p. 2884-94.
252. Ringe, R.P., et al., *Influences on the Design and Purification of Soluble, Recombinant Native-Like HIV-1 Envelope Glycoprotein Trimers*. J Virol, 2015. **89**(23): p. 12189-210.
253. Go, E.P., D. Hua, and H. Desaire, *Glycosylation and disulfide bond analysis of transiently and stably expressed clade C HIV-1 gp140 trimers in 293T cells identifies disulfide heterogeneity present in both proteins and differences in O-linked glycosylation*. J Proteome Res, 2014. **13**(9): p. 4012-27.
254. Ringe, R.P., et al., *Cleavage strongly influences whether soluble HIV-1 envelope glycoprotein trimers adopt a native-like conformation*. Proc Natl Acad Sci U S A, 2013. **110**(45): p. 18256-61.
255. Navis, M., et al., *HIV-1 receptor binding site-directed antibodies using a VHI-2 gene segment orthologue are activated by Env trimer immunization*. PLoS Pathog, 2014. **10**(8): p. e1004337.
256. Martinez, P., et al., *Primate immune responses to HIV-1 Env formulated in the saponin-based adjuvant AbISCO-100 in the presence or absence of TLR9 co-stimulation*. Sci Rep, 2015. **5**: p. 8925.
257. Guenaga, J., et al., *Structure-Guided Redesign Increases the Propensity of HIV Env To Generate Highly Stable Soluble Trimers*. J Virol, 2015. **90**(6): p. 2806-17.
258. Sarkar, A., et al., *Structure of a cleavage-independent HIV Env recapitulates the glycoprotein architecture of the native cleaved trimer*. Nat Commun, 2018. **9**(1): p. 1956.

259. Dubrovskaya, V., et al., *Targeted N-glycan deletion at the receptor-binding site retains HIV Env NFL trimer integrity and accelerates the elicited antibody response*. PLoS Pathog, 2017. **13**(9): p. e1006614.
260. Zhou, T., et al., *Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation*. Cell Rep, 2017. **19**(4): p. 719-732.
261. Cale, E.M., et al., *Virus-like Particles Identify an HIV V1V2 Apex-Binding Neutralizing Antibody that Lacks a Protruding Loop*. Immunity, 2017. **46**(5): p. 777-791 e10.
262. Ingale, J., et al., *High-Density Array of Well-Ordered HIV-1 Spikes on Synthetic Liposomal Nanoparticles Efficiently Activate B Cells*. Cell Rep, 2016. **15**(9): p. 1986-99.
263. Arnaout, R., et al., *High-resolution description of antibody heavy-chain repertoires in humans*. PLoS One, 2011. **6**(8): p. e22365.
264. Poulsen, T.R., et al., *Limits for antibody affinity maturation and repertoire diversification in hypervaccinated humans*. J Immunol, 2011. **187**(8): p. 4229-35.
265. Scheid, J.F., et al., *Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals*. Nature, 2009. **458**(7238): p. 636-40.
266. Kepler, T.B., *Reconstructing a B-cell clonal lineage. I. Statistical inference of unobserved ancestors*. F1000Res, 2013. **2**: p. 103.
267. Meng, W., et al., *An atlas of B-cell clonal distribution in the human body*. Nat Biotechnol, 2017. **35**(9): p. 879-884.